universidade de aveiro

U.PORTO
FLUP FACULDADE DE LETRAS
UNIVERSIDADE DO PORTO

3º CICLO

INFORMAÇÃO E COMUNICAÇÃO EM PLATAFORMAS DIGITAIS

Os Repositórios Institucionais das Universidades Federais do Brasil: Um Modelo de Política de Preservação Digital

Laerte Pereira da Silva Júnior

D

2017

Laerte Pereira da Silva Júnior

Os Repositórios Institucionais das Universidades Federais do Brasil: Um Modelo de Política de Preservação Digital

Tese realizada no âmbito do Doutoramento em Informação e Comunicação em Plataformas Digitais, orientada pelo professor-doutor Armando Malheiro da Silva, professor associado com agregação da Faculdade de Letras da Universidade do Porto e coorientada pela professora-doutora Maria Manuel Borges, professora auxiliar da Faculdade de Letras da Universidade de Coimbra.

Faculdade de Letras da Universidade do Porto

Julho de 2017

Os Repositórios Institucionais das Universidades Federais do Brasil: Um Modelo de Política de Preservação Digital

Laerte Pereira da Silva Júnior

Tese realizada no âmbito do Doutoramento em Informação e Comunicação em Plataformas Digitais, orientada pelo professor-doutor Armando Malheiro da Silva, professor associado com agregação da Faculdade de Letras da Universidade do Porto e coorientada pela professora-doutora Maria Manuel Borges, professora auxiliar da Faculdade de Letras da Universidade de Coimbra.

Membros do Júri

Professora-doutora Maria da Graça Lisboa Castro Pinto
Presidente – Faculdade de Letras da Universidade do Porto

Professora-doutora Ana Alice Rodrigues Pereira Baptista
Vogal – Escola de Engenharia da Universidade do Minho

Professora-doutora Maria Cristina Vieira de Freitas
Vogal – Faculdade de Letras da Universidade de Coimbra

Professora-doutora Lídia de Jesus Oliveira Loureiro da Silva
Vogal – Departamento de Comunicação e Arte da Universidade de Aveiro

Professora-doutora Maria Elisa Ramos de Morais Cerveira
Vogal – Faculdade de Letras da Universidade do Porto

Professora-doutora Maria Manuel Lopes de Figueiredo Costa Marques Borges
Vogal – Faculdade de Letras da Universidade de Coimbra

"It is not possible to have access without preservation."

National Library of Wales.

Sumário

Agradecimentos

A meus pais, *Diana* e *Laerte*, pelo amor com que me acolhem sempre e pelo apoio que me deram para estudar em outro país. A meu pai, revisor da linguagem deste trabalho e um doutor *ipso facto* da língua portuguesa, toda a minha reverência.

A meu filho *Jonas* pela alegria com que celebra as minhas conquistas acadêmicas e pela paciência em suportar a minha ausência nos últimos três anos.

A minhas irmãs, *Tânia* e *Telma,* pela preocupação que têm em me ver bem, em todos os sentidos, e a meu irmão *Julinho* por me divertir com suas criativas estórias de humor. A alegria de tê-los em minha vida é indizível.

A meus sobrinhos *Victor*, *Lindemberg*, *Gabriel*, *Rafael*, *Romeu*, *Fabinho* e a minhas sobrinhas *Daniela*, *Juliana* e *Larissa*. Vocês são uma fonte de inspiração.

A minha esposa *Thais* pelo incentivo de pleitear uma bolsa da CAPES, por ter escolhido a Universidade do Porto para estudarmos, pelas discussões que me iluminaram, pela paciência em me ouvir, pela força nos momentos de estresse, pelo carinho. Eu não teria realizado esse doutoramento sem você.

A meus amigos *Mana* e *Matheus* (meu *brother*), pela força de sempre, pela alegria da nossa convivência.

À família *Gallotti* pela amizade, carinho e solidariedade em nossa caminhada no Porto.

A meu orientador, professor-doutor *Armando Malheiro da Silva*, por ter aceitado a colaboração da professora-doutora *Maria Manuel Borges* para desenvolvermos a tese. Sua disposição para o debate norteou essa caminhada.

À professora-doutora *Maria Manuel Borges* pelo inestimável ensinamento sobre o papel dos repositórios institucionais na comunicação científica em acesso aberto e pela incansável disposição para esclarecer dúvidas, com sua peculiar objetividade.

Ao júri pelo desenvolvimento das provas de doutoramento e pelo tempo destinado à apreciação do trabalho de investigação (nomeadamente, à professora-doutora *Maria da Graça Lisboa Castro Pinto*, professora-doutora *Ana Alice Rodrigues Pereira Baptista*, professora-doutora *Maria Cristina Vieira de Freitas*, professora-doutora *Lídia de Jesus*

Oliveira Loureiro da Silva, professora-doutora *Maria Elisa Ramos de Morais Cerveira* e à professora-doutora *Maria Manuel Lopes de Figueiredo Costa Marques Borges*).

Aos professores das disciplinas cursadas neste Programa por me incentivarem com paciência e por ouvirem as minhas indagações, especialmente por propiciarem um exercício de livre pensamento nos debates ocorridos em sala de aula.

A todos os meus colegas de doutoramento pela convivência agradável e partilha de ideias sobre nossos próprios trabalhos.

Ao doutor *Miguel Ángel Márdero Arellano*, coordenador da Rede Brasileira de Serviços de Preservação Digital e pesquisador do IBICT, por despertar o meu interesse em pesquisar na área de política de preservação digital e por revisar o meu projeto para o ingresso neste programa doutoral.

A *Trevor Owens*, arquivista da *Library of Congress*, pelos vários *e-mails* nos quais discutimos os níveis de preservação digital modelados pela NDSA.

Ao professor-mestre *Daniel W. Noonan* pelas discussões sobre o *framework* de política de preservação digital da *Ohio State University Libraries.*

À professora-mestre *Andreia Silva Almeida* pelos esclarecimentos de dúvidas nas interpretações do referencial teórico na língua inglesa.

Às administradoras do repositório da UFRN pelo auxílio no pré-teste do questionário da pesquisa.

A meu colega *Ariosvaldo Borges Patrício* por ter aceitado me substituir na Assessoria de Informática durante o meu período de afastamento para cursar o doutorado.

Ao Conselho do Centro de Ciências Humanas, Letras e Artes da UFPB por aprovar o meu afastamento para cursar o doutorado na Universidade de Porto.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela concessão da minha bolsa de estudos na modalidade de Doutorado Pleno no Exterior.

Aos respondentes e todos quantos me ajudaram na produção desta pesquisa.

Resumo

O desenvolvimento da microeletrônica, das telecomunicações, da optoeletrônica e dos computadores, realizado ao longo da guerra fria, impulsionou a revolução da tecnologia da informação. A partir da década de 1970, entra em curso um movimento econômico e cultural, conduzido por governos, empresas e sociedade civil, que resulta no surgimento da Internet, o meio tecnológico mais transformador da Era da Informação. O novo paradigma tecnológico alterou o modo de organizar, acessar e produzir a ciência. Essas alterações transformaram o ciclo da comunicação científica resultando no que Michael Nentwich denomina de ciberciência. Os repositórios institucionais constituem um dos elementos da ciberciência. Garantir que eles sobrevivem à passagem do tempo exige a definição de uma política de informação para dar conta de todos os aspectos necessários ao desempenho do papel a que se destinam. Assim, o planejamento de uma política de preservação digital permite estabelecer ações de investimento na qualificação do pessoal tecnoadministrativo e na infraestrutura, bem como a adaptação de práticas de preservação digital consensuais no âmbito internacional. Ao extrapolarem-se as atividades meramente técnicas inerentes às estratégias de preservação digital, pode-se garantir o acesso a longo prazo do acervo dos repositórios institucionais. Esta tese tem por objetivo geral o de propor um modelo de política de preservação digital para os repositórios das universidades federais brasileiras. Adotamos um método misto com alcance exploratório e descritivo. O universo da pesquisa é constituído por 38 universidades. As fontes de coleta de dados foram o OpenDOAR, as *homepages* dos repositórios e um questionário *online* aplicado aos administradores. A análise dos dados revelou a ausência de uma política de preservação digital e as características da cultura de preservação digital nas universidades federais. Diante desse resultado, propusemos um modelo de política de preservação digital, com base nos atributos de um repositório digital confiável, adequando-os à realidade das universidades pesquisadas.

Palavras-chave: Repositório Institucional, Política de Acesso aberto, Modelo de Política de Preservação Digital.

Abstract

Development of microelectronics, telecommunications, optoelectronics and computers, achieved throughout the Cold War, fueled the revolution of information technology. From the 1970s onwards, a technological, economic and cultural movement arises, led by governments, companies and the civil society, which results in the emergence of the Internet, the most transformative technological means of the Information Age. The new technological paradigm changed the way of organizing, accessing, and producing science. These changes transformed the scientific communication cycle, resulting in what Michael Nentwich denominates as cyberscience. Institutional repositories are one of the technological elements of cyberscience. Survival over the years imposes the definition of an information policy which aims at dealing with all aspects necessary for the fulfillment of the role they are meant to play. A digital preservation policy planning enables the establishment of investment actions related to the qualification of the staff and to the improvement of the infrastructure. Furthermore, it enables the adaptation of digital preservation practices agreed internationally. By extrapolating the purely technical activities inherent to the digital preservation strategies, one can ensure long-term access to the institutional repository collections. Thus, this thesis has as general objective the proposal of a policy model for the digital preservation in institutional repositories of the brazilian federal universities. We adopted a mixed method with exploratory and descriptive outreach. The research universe is made up of 38 universities. The sources of data collection were the OpenDOAR, the homepages of the repositories and an online questionnaire applied to administrators. Data analysis revealed the absence of a digital preservation policy and the characteristics of the digital preservation culture in federal universities. Given this result, we proposed a digital preservation policy model, based on the attributes of a trusted digital repository, adapting them to the reality of the researched universities.

Keywords: Institutional Repository, Open Access Policy, Digital Preservation Policy Model.

Índice de figuras

Índice de gráficos

Índice de quadros

Lista de abreviaturas e siglas

AIP	*Archival Information Package*
DARPA	*Defense Advanced Research Projects Agency*
FINEP	Financiadora de Estudos e Projetos
FURG	Universidade Federal do Rio Grande
IBICT	Instituto Brasileiro de Informação em Ciência e Tecnologia
ICPSR	*Inter-university Consortium for Political and Social Research*
JISC	*Joint Information Systems Committee*
METS	*Metadata Encoding & Transmission Standard*
NDIIP	*National Digital Information Infrastructure and Preservation Program*
NDSA	*National Digital Stewardship Alliance*
OAIS	*Open Archival Information System*
OpenDOAR	*Directory of Open Access Repositories*
PDI	Plano de Desenvolvimento Institucional
PII	Política de Informação Institucional
PPD	Política de Preservação Digital
PREMIS	*Preservation Metadata : Implementation Strategies*
RCAAP	Repositório Científico de Acesso aberto de Portugal
RI	Repositório Institucional
RSP	*Repository Supported Project*
SCAPE	*Scalable Preservation Environments*
TIC	Tecnologias de Informação e Comunicação
TRAC	*Trustworthy Repositories Audit & Certification: Criteria and Checklist*
UFAL	Universidade Federal de Alagoas

UFBA	Universidade Federal da Bahia
UFC	Universidade Federal do Ceará
UFES	Universidade Federal do Espírito Santo
UFF	Universidade Federal Fluminense
UFG	Universidade Federal de Goiás
UFGD	Universidade Federal da Grande Dourados
UFJF	Universidade Federal de Juiz de Fora
UFLA	Universidade Federal de Lavras
UFMA	Universidade Federal do Maranhão
UFMG	Universidade Federal de Minas Gerais
UFMS	Universidade Federal do Mato Grosso do Sul
UFMT	Universidade Federal de Mato Grosso
UFOP	Universidade Federal de Ouro Preto
UFPA	Universidade Federal do Pará
UFPB	Universidade Federal da Paraíba
UFPE	Universidade Federal de Pernambuco
UFPEL	Universidade Federal de Pelotas
UFPI	Universidade Federal do Piauí
UFPR	Universidade Federal do Paraná
UFRB	Universidade Federal do Recôncavo da Bahia
UFRGS	Universidade Federal do Rio Grande do Sul
UFRN	Universidade Federal do Rio Grande do Norte

UFS	Universidade Federal de Sergipe
UFSC	Universidade Federal de Santa Catarina
UFSM	Universidade Federal de Santa Maria
UFT	Universidade Federal do Tocantins
UFTPR	Universidade Tecnológica Federal do Paraná
UFU	Universidade Federal de Uberlândia
UFV	Universidade Federal de Viçosa
UFVJM	Universidade Federal dos Vales do Jequitinhonha e Mucuri
UNB	Universidade de Brasília
UNIFEI	Universidade Federal de Itajubá
UNIFESP	Universidade Federal de São Paulo
UNILA	Universidade Federal da Integração Latino-Americana
UNIPAMPA	Universidade Federal do Pampa
UNIR	Universidade Federal de Rondônia
URL	*Uniform Resource Locator*

Introdução

No período da guerra fria, as tecnologias de informação começaram a ser desenvolvidas por grandes projetos governamentais, principalmente os dos Estados Unidos da América, e desdobraram-se em projetos de interesse para a sociedade civil, transformando-lhe o modo de vida, com ressonância na economia e nas comunicações de modo global. As tecnologias da informação fizeram surgir um novo paradigma com características sociotécnicas sem precedentes na história das civilizações. O resultado mais proeminente da interação entre a sociedade, o Estado e a tecnologia informacional é a Internet, considerado o meio tecnológico mais revolucionário da Era da Informação (Castells, 1999).

Na academia, as tecnologias de informação e comunicação (TICs) começaram a ser difundidas no início da década de 1980, com o crescente uso de computadores em rede, Internet, *e-mail*, lista de discussão, bases de dados *online* e *off-line*, buscadores, publicações e conferências eletrônicas, etc. Essa cultura caracteriza um novo modo do fazer científico denominado de ciberciência, que consiste em todas as atividades acadêmicas e de pesquisa geradas no ciberespaço. O modelo de impacto das TICs na academia, elaborado por Nentwich (2003 a), busca demonstrar como elas mudam o sistema de comunicação acadêmica tradicional para o da ciberciência, isto é, os efeitos que essas mudanças provocam em toda a comunicação estruturada da academia, no modo de trabalho da comunidade e na representação do conhecimento. Um fator determinante na evolução de uma tecnologia é a pressão de demanda do grupo social caracterizado como seus usuários. Por essa razão, o modelo de impacto ajuda a entender a influência dos fatores institucionais, funcionais e relacionados com a comunidade acadêmica na passagem da fase inicial da *web* para a denominada *web* 2.0.

Nesse novo ambiente *web*, foram produzidas as formas mais céleres de comunicação e disseminação do conhecimento e testaram-se novas modalidades de expressão que abrangem todos os pontos do ciclo da comunicação da ciência. O acesso aberto ao conhecimento científico, conforme foi enunciado na declaração da *Budapeste Open Access Initiative* (BOAI, 2002), se juntou a um processo tradicional de democratização da comunicação dos resultados de pesquisas. A convergência desses dois

modos de difundir a ciência, com o apoio da *internet*, revolucionou a forma de acesso e distribuição da literatura científica:

> Uma antiga tradição e uma nova tecnologia convergiram para tornar possível um avanço histórico. A antiga tradição é a disposição de cientistas e acadêmicos em publicar o fruto de suas pesquisas sem remuneração, em nome da transparência e democratização do conhecimento. A nova tecnologia é a internet. O avanço histórico que eles possibilitam é a distribuição da literatura acadêmica arbitrada por toda a extensão do globo e o acesso totalmente irrestrito e gratuito por parte de qualquer cientista, acadêmico, professor, estudante ou outro interessado (BOAI, 2002).

Na declaração da BOAI (2002) definem-se as duas vias de acesso aberto ao conhecimento científico: a via verde – a via dos repositórios institucionais (RIs) – e a via dourada – a via dos periódicos em acesso aberto.

Um RI é um tipo de repositório digital, predominantemente de natureza acadêmica, que coleta, dissemina e preserva a produção intelectual de uma instituição, constituindo-se em uma ferramenta para a inserção da comunidade institucional no movimento *Open Access* (Burns, Lana e Budd, 2013; SDUM, 2015). Um RI, quando permite o acesso aberto ao seu conteúdo e ao *harvesting* dos seus metadados, é tipificado como um repositório de acesso aberto (Heery e Anderson, 2005).

A implementação de um RI próprio é bastante simples e de baixo custo, pois diversas plataformas robustas utilizadas no mundo todo estão disponíveis sob licença *free software* ou *open source*. Por exemplo: o *EPrints* é distribuído sob a licença GPLv3/LGPLv3 e o *DSpace* é distribuído sob a *BSD open source license*. Todavia, a cultura institucional é a principal barreira a ser superada com vistas ao sucesso do funcionamento de um sistema desse tipo, pois é preciso que se desenvolva um processo de estímulo e conscientização da comunidade de produtores dos objetos digitais que serão armazenados. Para tanto, Borges (2006) recomenda que seja priorizado o estabelecimento das condições de operacionalidade do repositório em três componentes: o cultural, o organizacional e o técnico. A análise dessas condições revela que o desenvolvimento de uma política institucional e a criação de regras de utilização são peças fundamentais no processo de constituição de um RI. De fato, uma política institucional é estabelecida por meio de portarias, resoluções e demais instrumentos legais, o que compromete a instituição e a comunidade universitária, com relação aos direitos e deveres de ambas as partes. Naturalmente, haverá políticas diferenciadas, como a que estabelece regras para o depósito e para os direitos de autor, previstas no componente organizacional, assim como

regras relacionadas com os formatos de arquivo, metadados e as estratégias de preservação, previstas no componente técnico. Esses dois últimos componentes são fundamentais para se construir uma política de preservação digital.

A preservação digital é o conjunto de ações e intervenções requeridas para garantir o acesso (contínuo e confiável) aos objetos digitais autênticos, ao longo do tempo em que forem considerados válidos. Isso abrange tanto as atividades técnicas quanto as questões estratégicas e organizacionais que implicam a sobrevivência e o gerenciamento do material digital (Pennock, 2006). Todos esses aspectos precisam ser coordenados por meio de um mandato expressado em uma política de preservação digital, a qual é definida pela norma ISO 16363 como uma declaração escrita, promulgada por uma instituição que descreve a abordagem a ser considerada para a preservação dos objetos armazenados. Diversos *frameworks* para esse tipo de política estão disponíveis na *Web*, em portais de projetos de preservação digital ou de instituições preocupadas em compartilhar seus modelos teóricos. Por exemplo: *Joint Information Systems Committee*; *Canadian Heritage Information Network*[1]; *The Inter-university Consortium for Political and Social Research.* Todavia, o Catálogo de Elementos de Política de Preservação – *Catalogue of Preservation Policy Elements* (Sierman *et al*., 2014) – diferencia-se dos demais *frameworks* disponíveis, porque propicia uma visão geral dos elementos essenciais que uma instituição precisará formular para correlacionar os níveis do seu conjunto de políticas de preservação, assim estabelecidos: políticas de orientação, políticas de procedimentos de preservação e políticas de controle. Esses três níveis são interligados e identificados na sua adequação a um determinado propósito. Dessa forma, uma instituição poderá criar uma política de preservação de forma mais objetiva e estará mais preparada para automatizar essa política. O nível de políticas de orientação é composto por dez elementos. Cada um deles possui elementos de procedimentos de preservação com sua respectiva política de controle. As políticas de orientação são denominadas como se seguem: autenticidade, preservação de *bit*, preservação funcional, objeto digital, metadados; direitos, padrões, acesso, aspectos organizacionais, auditoria e certificação. Todos eles serão devidamente abordados neste trabalho.

[1] Disponível na WWW em: http://canada.pch.gc.ca/eng/1445528225711

As universidades federais do Brasil têm implantado seus repositórios com o apoio de uma parceria entre o Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) e a empresa pública Financiadora de Estudos e Projetos (FINEP) ou com a iniciativa voluntária de membros da comunidade acadêmica. Segundo Ribeiro (2012), o projeto IBICT/FINEP contemplou 40 instituições de ensino e pesquisa, as quais se comprometeram com a elaboração de uma política de informação para cuidar do funcionamento, depósito, preservação digital, acesso e comunicação de suas respectivas produções científicas. Entretanto, em sua pesquisa, a autora constatou que os RIs estão sendo criados sem uma política de preservação digital. No caso dos RIs criados por iniciativas voluntárias, podemos citar como exemplo o caso do RI da Universidade Federal da Paraíba (UFPB). Esse repositório foi criado sem qualquer vínculo com o projeto coordenado pelo IBICT, contando-se apenas com o esforço de bibliotecários, técnicos e professores interessados na temática.

Dentre as trinta e oito universidades que possuem um RI, apenas dezesseis adotaram o *framework* do IBICT[2] para criar suas políticas de informação institucional (PIIs). Por meio de uma PII elas manifestam a intenção de preservar o material digital armazenado em seus respectivos RIs, mas não registram qualquer política dessa natureza, seja nos próprios RIs, seja no *OpenDOAR*, onde estão cadastrados quatorze dos RIs que adotaram o modelo de PII do IBICT. Em casos como o da UFPB, onde nem sequer o RI está institucionalizado por meio de uma PII, o caminho para a construção de uma política de preservação digital[3] precisa ser pavimentado pela assimilação de uma cultura de acesso aberto, para dar sentido à existência de um RI. Conforme Marcondes e Sayão (2009, p. 9), nesta cultura um RI desempenha o papel de um "ator político" da comunicação científica. De todo modo, tanto nas universidades que criaram sua PII quanto naquelas que não criaram esse tipo de política, não há registro de uma política específica para a preservação digital. Diante desta realidade, considerando nosso interesse de obter um

[2] O *framework* para uma PII, proposta pelo IBICT, está descrito no trabalho de Kuramoto (2009).

[3] Uma política de preservação digital é uma declaração escrita, autorizada pela instituição mantenedora de um repositório, que descreve as abordagens que serão implementadas pelo repositório para a preservação dos objetos armazenados (CCSDS, 2012). Uma política de preservação é consistente com uma Política Institucional de Informação.

conhecimento mais além do mero *know-how* técnico, propomo-nos a pesquisar a problemática da política de preservação digital, a partir do seguinte questionamento:

Quais são as iniciativas adotadas pelas universidades federais do Brasil que podem contribuir para a definição de um modelo de política de preservação digital em seus repositórios institucionais?

No intuito de encontrarmos a resposta, elaboramos os seguintes objetivos de trabalho:

Objetivo geral: propor um modelo de *framework* de política de preservação digital para os repositórios institucionais das universidades federais brasileiras.

Objetivos específicos:

- Analisar as grandes mudanças na academia com a emergência da ciberciência;
- Caracterizar a origem e a função dos RIs;
- Analisar os modelos de *framework* de política de preservação digital elaborados por iniciativas norte-americanas e europeias;
- Verificar a existência de uma política de preservação digital para os RIs das universidades federais brasileiras.
- Apurar a percepção dos administradores dos RIs das universidades federais sobre as questões de preservação digital.

O universo da pesquisa é formado por 38 universidades federais e todas elas contribuíram com a coleta de dados por meio de um questionário online. A pesquisa é composta pelas seguintes etapas: (i) exploração da literatura para se estabelecerem os elementos estruturantes de uma política de preservação digital; (ii) análise dos modelos de *framework* desse tipo de política, na América do Norte e Europa; (iii) explorar o registro dos RIs das universidades federais brasileiras no OpenDOAR, particularmente na seção onde deveria estar descrita a política de preservação dos repositórios; (iv) análise das PIIs em busca de uma indicação da existência de uma política de preservação digital; (v) aplicação de um questionário online direcionado aos administradores dos repositórios das universidades federais; (vi) análise e discussão dos resultados obtidos, (vii) formulação de um modelo de *framework* de política de preservação para os RIs.

Este trabalho está estruturado em 6 capítulos. O capítulo 1 discorre sobre a evolução histórica das TICs e sua penetração na academia para instrumentalizar um novo modo de comunicação acadêmica e fazer científico denominado de ciberciência. Nesse contexto, dá-se a crise dos periódicos, um fator determinante para o surgimento do movimento de acesso aberto. Tal movimento tem nos RIs o pavimento da sua via verde. O capítulo 2 trata dos conceitos teóricos fundamentais para uma política de preservação digital: autenticidade; preservação de *bit;* preservação funcional; objeto digital; metadados; direitos; padrões; acesso; aspectos organizacionais; auditoria e certificação. O capítulo 3 analisa os modelos de *frameworks* para uma política de preservação em repositórios digitais, produzidos por projetos da América do Norte e Europa, com destaque para o OpenDOAR, o modelo do JISC, o catálogo de política do SCAPE e o modelo do InterPARES. O capítulo 4 descreve a metodologia utilizada para esta pesquisa, cujos métodos são o quantitativo e o qualitativo, com o alcance de um estudo exploratório e descritivo, conforme conceituação dada por Sampieri, Collado e Lucio (2013). O capítulo 5 desenvolve uma análise das Políticas de Informação Institucional das universidades federais brasileiras e das características da cultura de preservação digital nessas instituições. Essa análise revelou a ausência de uma política de preservação digital. Por fim, o capítulo 6 propõe um modelo de *framework* de política de preservação digital para as universidades federais do Brasil aplicarem a um programa de preservação digital do acervo dos seus RIs.

Capítulo 1 – O papel das tecnologias de informação e comunicação na constituição dos repositórios institucionais de acesso aberto

O desenvolvimento das tecnologias de informação iniciado, no período da guerra fria, por grandes projetos governamentais, especialmente o do governo dos Estados Unidos, derivou para os interesses da sociedade civil transformando-lhe o estilo de vida, com impacto na economia e nas comunicações em escala global. As tecnologias de informação trouxeram um novo paradigma, com características sociais e técnicas sem precedentes na história da humanidade.

Segundo Castells (1999), a revolução das tecnologias de informação revela a complexidade de uma nova economia, sociedade e cultura em formação. Entretanto, a tecnologia não estabelece uma sociedade nem tampouco a sociedade traça o percurso da modificação tecnológica, pois inúmeros fatores, dentre eles a criatividade e o empreendedorismo, interferem no desenvolvimento da descoberta científica, inovação tecnológica e aplicações para a sociedade, de modo que o resultado último depende de um padrão de interatividade complexo. Contrapõem-se ao determinismo tecnológico o fato de a tecnologia ser a própria sociedade e esta não poder ser entendida ou representada sem o seu ferramental tecnológico. Por exemplo: na década de 1970, surgiu um novo paradigma tecnológico, ordenado com base na tecnologia da informação. Simultaneamente, um dado segmento da sociedade norte-americana, interagindo com o mundo tanto no aspecto econômico quanto geopolítico, estabeleceu um novo modo de produção, comunicação, gestão e vida.

Como a consolidação desse paradigma ocorreu nos Estados Unidos, na Califórnia da década de 1970, é possível que isto seja a causa do amplo desenvolvimento das novas tecnologias de informação. Por exemplo: apesar de o desenvolvimento da indústria eletrônica ter sido impulsionado pelo financiamento militar e empresarial na década de 1940 e na de 1960, o grande salto tecnológico da década de 1970, provavelmente, está relacionado com a cultura da liberdade, inovação e estímulo empreendedor cultivado nos *campi* norte-americanos da década de 1960. Assim, procura-se enfatizar a criação de dispositivos personalizados, a interatividade, a formação de redes e a busca interminável

de novas descobertas tecnológicas, ainda que não houvesse aplicação comercial. De forma meio inconsciente, a revolução da tecnologia difundiu o espírito libertário da década de 1960 pela cultura de maior significado em nossas sociedades. Entretanto, as novas tecnologias de informação foram disseminadas e apropriadas por vários países, culturas e organizações com diferenciados objetivos, o que resultou em uma aceleração das transformações tecnológicas, ampliando seus objetivos e diversificando suas fontes.

O entendimento da importância dos impactos sociais involuntários da tecnologia pode ser mais bem entendido com o seguinte exemplo:

A Internet foi uma criação dos estrategistas da *Defense Advanced Research Projects Agency* (DARPA) dos Estados Unidos para se prevenir contra um possível ataque nuclear soviético ao sistema de comunicação norte-americano. Com o passar do tempo a ARPANET, rede criada pela DARPA, transformou-se na base de uma rede horizontal de alcance global formada por milhares de redes de computadores. Essa rede foi apropriada por pessoas e grupos por todo o mundo e com os mais variados objetivos, completamente diferentes das preocupações de uma guerra fria extinta.

A sociedade, apesar de não determinar a tecnologia, pode sufocar o desenvolvimento tecnológico, principalmente por meio do poder do Estado. A história da tecnologia chinesa e da japonesa ensina que, para se entender a relação entre a tecnologia e a sociedade, faz-se necessário saber que o papel do Estado na inovação tecnológica é crucial em todo o processo, à proporção que manifesta e ordena as forças sociais majoritárias em um local e período determinados. A tecnologia demonstra a capacidade de uma sociedade em promover seu domínio tecnológico por meio das instituições sociais e do próprio Estado. O contexto histórico em que ocorre o desenvolvimento de forças produtivas evidencia as características tecnológicas e seus amálgamas com as relações sociais. Portanto, não há diferença entre a história da tecnologia chinesa e a da japonesa, como é o caso da revolução tecnológica da atualidade. Ela teve origem e foi difundida em uma época de reestruturação global do sistema capitalista, para o qual serviu como ferramenta de apoio. Por conseguinte, a nova sociedade que emerge do processo de reestruturação global é capitalista e informacional, conquanto apresente diferentes aspectos históricos nos vários países, consoante sua história, cultura, instituições e relação particular com o capitalismo global e com a tecnologia informacional (Castells, 1999).

As principais descobertas da eletrônica ocorreram durante a Segunda Guerra Mundial, quando foram inventados o primeiro computador programável e o transistor, base da microeletrônica e centro da revolução da tecnologia da informação do século XX. Entretanto, apenas na década de 1970, as novas tecnologias da informação foram largamente difundidas, em um movimento de aceleração do seu desenvolvimento cooperativo e consequente convergência em um novo paradigma sociotécnico, cujos aspectos centrais representam a base material da sociedade de informação. Assim, o novo paradigma da tecnologia da informação possui cinco características:

1) A informação é sua matéria prima: as tecnologias são criadas para agir sobre a informação, principalmente.
2) A penetrabilidade dos efeitos das novas tecnologias: o novo meio tecnológico molda a vida individual e coletiva, embora não a determine.
3) A lógica de redes em qualquer sistema ou conjunto de relações, que utilize as novas tecnologias de informação: estrutura o não estruturado sem comprometer a flexibilidade, porque o não estruturado é a alavanca da inventividade humana.
4) A flexibilidade: capacidade de reconfigurar os processos, as organizações e as instituições, sem que isso signifique uma destruição de qualquer desses elementos.
5) A crescente convergência de tecnologias específicas para um sistema altamente integrado. A microeletrônica, as telecomunicações, a optoeletrônica e os computadores formam um conjunto de campos que integram os sistemas de informação. A convergência desses campos na área da comunicação interativa, por exemplo, conduziram ao surgimento da Internet, provavelmente o meio tecnológico mais revolucionário da Era da Informação (Castells, 1999).

1.1 − Ciberciência: a transformação do fazer científico

Nentwich (2003 a) assegura que, desde o princípio da década de 1980, a comunidade acadêmica tem experimentado um considerável aumento no uso das tecnologias de informação e comunicação (TIC). Computadores em rede, *e-mail*, a

Internet, base de dados *online* e *offline*, a *World Wide Web*, publicações eletrônicas, lista de discussão e *newsgroups*, conferências eletrônicas, bibliotecas digitais e buscadores são apenas alguns dos exemplos que vêm influenciando o dia a dia da comunidade científica, desde aquela época. A partir dessa constatação, o autor propõe uma noção de ciberciência em oposição a ciência e a pesquisa tradicionais, isto é, feitas sem redes de computadores. Assim, a ciberciência designa o uso desses serviços e aplicações baseadas nas TICs para fins científicos. Para o propósito do seu estudo, Nentwich (2003 b, p. 22) define-a: "(…) as all scholarly and scientific research activities in the virtual space generated by the networked computers and by advanced information and communication technologies, in general."

Nentwich (2003 a) corrobora o conceito de revolução aplicado por Harnad (1991) ao desenvolvimento da comunicação e publicação eletrônica, porque ele impulsiona todas as atividades acadêmicas, potencializa a comunicação científica e eleva a qualidade das pesquisas.

As TICs causam alterações em todos os aspectos da vida acadêmica. Já interpuseram sua marca nas atividades de coletar, discutir, analisar e distribuir a informação científica. A mola propulsora dessas ações são as redes de computadores. Tal constatação conduziu a uma categorização dos tipos de atividade acadêmica no contexto do ciberespaço. São eles:

a) Produção do conhecimento (*knowledge production*) – Inclui a coleta de informação, produção, análise e gerenciamento de dados.
b) Processamento do conhecimento (*knowledge processing*) – Está contido na comunicação científica e é desenvolvido por meio da representação do conhecimento, da discussão, da avaliação e da cooperação.
c) Distribuição do conhecimento (*knowledge distribution*) – Publicação, implementação e ensino;
d) Ambiente institucional (*institutional setting*) – Equipamentos técnicos requeridos pelas atividades acadêmicas e estrutura organizacional.

Essas quatro categorias estão representadas no diagrama da figura 1.

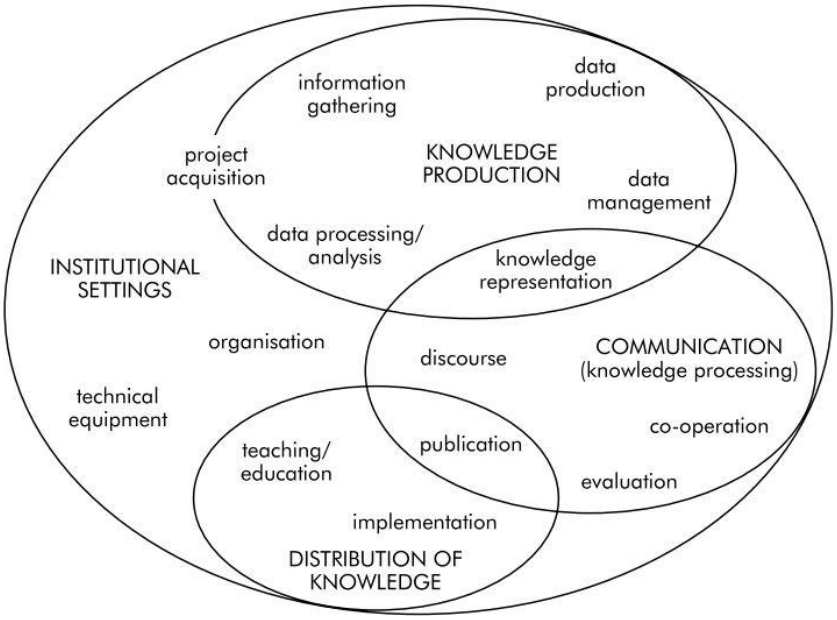

Figura 1 – Categorias de atividades acadêmicas e seus elementos estruturantes

Fonte: Nentwich, 2003 b, p. 24.

Na figura 1, nota-se que existem pontos de intersecção entre as categorias, tais como o que diz respeito à publicação (*publication*), representação do conhecimento (*knowledge representation*) e aquisição de projeto (*project acquisiton*). Esse diagrama é a base para se compararem as ferramentas de que um cientista tradicional e um cibercientista dispõem nestas categorias. De fato, em todas as categorias elencadas se podem testemunhar alterações cada vez mais significativas. Não são apenas a produção de conhecimento e a sua comunicação que são afetadas pelas TICs, mas sobretudo as modalidades onde a colaboração é a pedra de toque. O quadro 1 serve para mostrar que a ciberciência está afetando todas essas categorias em um processo de transformação tecnorganizacional.

Quadro 1 - Transformações do fazer científico na trilha do ciberespaço

		Technical-organisational transformation "Traditional" science → Cyberscience			
institutional set-up	Organisation	Traditional institute; guest researchers		Telework	Virtual institute
	Technical equipment	Typewriter; telephone; library	Stand-alone PC; fax	Internet connection	Multimedia PC; access to data networks
knowledge production	Project acquisition	Face-to-face/by letter/ telephone negotiations		E-mail exchange	Electronic procurement
	Information gathering	Libraries; personal conversations	Offline databases	Online databases; link collections; discussion lists	Digital libraries; knowbots
	Data production	Interviews; experiments	Electron. text analysis; simulation/modelling	Internet surveys	Distributed computing; virtual reality
	Data management	Card files; lists	Hypertextual card files; databases		Networked card files; de-central databases
	Data processing/ analysis	"With paper and pencil"	Electron. data-processing; expert systems	Modelling; simulations	Artificial intelligence
	Knowledge representation	Linear texts	Electron. text-processing; databases	Multimedia; hypertexts	Hypertext-databases
communication (knowledge processing)	Co-operation	Letters; telephone; personal meetings	Exchange of electron. Manuscripts	E-mail; de-central assembly of databases; software sharing	Groupware
	Discourse evaluation	Conferences; seminars; conversations (pers./tel.)		E-mail; discussion lists; skywriting	Online conferences; internet chatting
distri-bution	Publication	Print media	Submission of electron. Manuscripts	Parallel publication in WWW; E-pre-prints	Pure E-publications; "net of knowledge"
	Teaching/education	Traditional teaching (seminars, lectures)	Correspondence courses	Multimedia manuals	Virtual university

Fonte - Nentwich (2003 b, p. 25)

Nentwich (2003 b) esclarece que o propósito dessa visualização é o de mostrar que em todas as linhas do quadro aparecem aplicações típicas do ciberespaço. Em outras palavras: as aplicações do ciberespaço surgem em todas as atividades acadêmicas e em todos os elementos que as estruturam. Por exemplo: a distribuição do conhecimento por meio do ensino em sala de aula, seminários e palestras ainda acontecem. Livros e cópias de artigos acadêmicos feitos a mão também desempenham um papel importante. Entretanto, o que se iniciou como um curso por correspondência transformou-se em cursos multimídias, em CD-ROM ou *online*, acessíveis na *Web*, muitos dos quais gratuitos, como os *Massive Open Online Courses* (MOOC). No gerenciamento de dados (*data management*), as bases que estavam em computadores *stand-alone*, agora estão acessíveis *online* e podem ser alimentadas, cooperativamente, por pesquisadores dispersos geograficamente. As publicações impressas; continuam a receber submissões dos manuscritos em versão eletrônica. Todavia, no decorrer do tempo, passou-se a uma

divulgação paralela na *Web* por meio de *preprints*, de modo que hoje em dia temos publicações totalmente eletrônicas contribuindo para a formação de uma rede de conhecimento (*net of knowledge*). Acrescente-se que tanto os *preprints* como os *postprints* podem estar depositados em repositórios institucionais ou temáticos – uma tecnologia que tem sido adotada largamente pelas universidades e centros de pesquisa do mundo todo.

Segundo Nentwich (2003 b), é óbvio que as TICs possuem o potencial de afetar quase todos os aspectos das atividades acadêmicas. Isso é o ponto de partida do seu estudo. A principal questão que ele propõe para ser investigada, ao longo de sua obra, deve procurar responder como as TICs causam impacto no mundo acadêmico. Para tanto, o referido autor elaborou um modelo básico para iniciar a descrição do fenômeno (figura 2).

A mudança no sistema de comunicação acadêmica é o ponto central do modelo na evolução do sistema de comunicação acadêmica tradicional para o da ciberciência. O impacto das TICs na academia em geral representa os efeitos que essas mudanças provocam na comunicação acadêmica e, por extensão, em toda a comunicação estruturada da academia. O impacto na substância da pesquisa diz respeito ao impacto causado à metodologia, ao modo do trabalho dos pesquisadores e à representação do conhecimento.

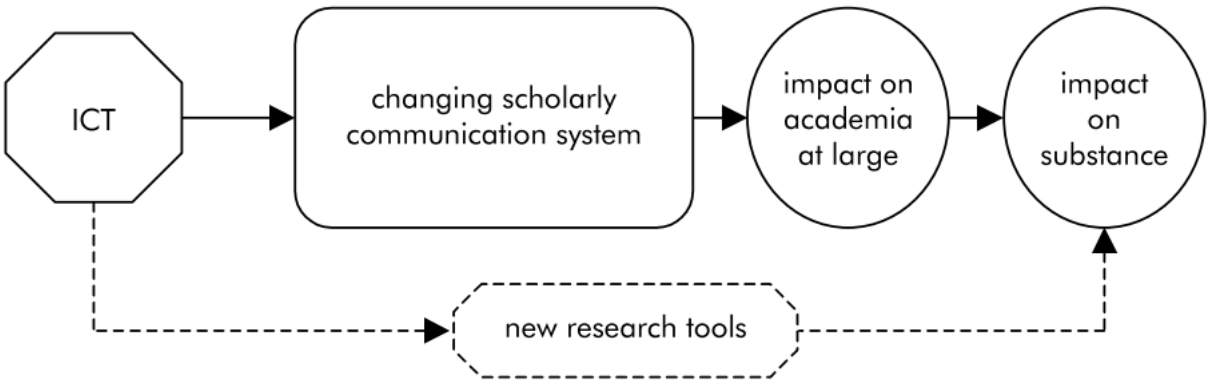

Figura 2 – Modelo básico do impacto das TICs na academia

Fonte: Nentwich (2003 b, p. 30)

O termo TIC envolve tanto a transmissão de dados, analógicos ou digitais, de pessoa para pessoa e pessoa para máquina, quanto as máquinas (*lato sensu*) que processam a informação em si mesmas (computadores, instrumentos, softwares, base de

dados). A ciberciência tem o foco num subconjunto das tecnologias da comunicação, *stricto sensu*, particularmente na Internet como a principal inovação tecnológica. No entanto, temos que considerar outros subconjuntos das TICs, como as *research tools*, que também influem nos resultados das atividades de pesquisa. O poder de afetação dos sistemas especializados, da inteligência artificial, da simulação e dos *softwares* processados em computadores *stand-alone* é um tema para um outro estudo, contudo, esses elementos foram incluídos apenas como tópicos, no conjunto das TICs que impactaram o fazer científico na academia (quadro 2).

Quadro 2 – Tecnologias de Informação e Comunicação

TIC

- Tecnologias de Informação
  (computadores, Inteligência Artificial, base de dados...)

- Tecnologias de Comunicação
  (redes de computadores)
  - Pessoa para pessoa
  - Pessoa para máquina
  - Máquina para máquina

Fonte - Nentwich (2003 b, p. 31)

O que vai determinar a evolução de uma tecnologia e sua demanda de uso é o interesse do grupo social que a utiliza; por isso, o modelo de impacto analisa a influência dos fatores institucionais, funcionais e relacionados com os atores na passagem da primeira para a segunda geração das TICs na ciberciência.[4] Entre os diversos tipos de elemento tecnológico da ciberciência encontram-se os repositórios digitais, cuja tipologia de maior interesse para este trabalho diz respeito aos RIs.

[4] A primeira geração das TICs na ciberciência é composta por *e-mail*, bases bibliográficas, publicações eletrônicas, *homepages*, lista de *links* etc. A segunda é constituída por videoconferências, publicações multimídia, etc.

1.2 − O surgimento dos repositórios institucionais

A comunicação científica é fundamental para o progresso científico e tecnológico. Como demonstra a história da produção da literatura científica, esse tipo de publicação aumentou, vertiginosamente, durante a Segunda Guerra Mundial, o que levou a uma "explosão informacional", identificada por Vannevar Bush como um aumento exponencial da informação e de seus registros, especialmente em ciência e tecnologia. Dessa forma, surgiu a moderna indústria da informação e as concepções que a direcionam (Saracevic, 1996). É neste contexto que despontam as editoras científicas explorando o mercado de periódicos, a ponto de reprimir a demanda, devido aos altos custos das assinaturas das publicações:

> A aparente estabilidade de que gozava o sistema de comunicação científica mundial foi abalada quando estourou a chamada crise dos periódicos, em meados da década de 1980, que já vinha se anunciando desde a década de 70. O gatilho da crise foi a impossibilidade de as bibliotecas universitárias e de pesquisa americanas continuarem a manter suas coleções de periódicos e a corresponder a uma crescente demanda de seus usuários, impossibilidade decorrente da falta de financiamento para a conta apresentada pelas editoras, cada ano mais alta, mais alta mesmo que a inflação e outros índices que medem a economia. Isso já vinha acontecendo nos países em desenvolvimento, inclusive no Brasil, cujas bibliotecas já não conseguiam manter suas coleções atualizadas, mas a crise só detonou quando atingiu as universidades norte-americanas (Mueller, 2006, p. 31).

A crise dos periódicos estimulou o surgimento de um movimento em favor do acesso livre ao conhecimento científico. Assim, a partir do início da década de 1990, iniciam-se vários trabalhos para se resolver o impasse da difusão e acesso aos artigos científicos. Segundo Mueller (2006) e Borges (2006), o artigo *Scholarly Skywriting and the Prepublication Continuum of Scientific Inquiry* (Harnad, 1990) é um marco na história do acesso livre aos textos acadêmicos, assim como a criação do arXiv[5] por Paul Ginsparg, um repositório temático, na área da Física de partículas de alta energia. Além dos repositórios temáticos, começam a surgir os repositórios institucionais com o objetivo de refletir toda a produção intelectual de uma instituição:

> O sucesso do arXiv, um repositório disciplinar, foi logo seguido pelo lançamento de serviços similares para outras áreas temáticas e instituições de grande dimensão e, eventualmente confluiu, em 1999, no surgimento da *Open*

[5] Sobre o nascimento do arXiv veja, por exemplo: <http://arxiv.org/pdf/1108.2700.pdf>

> *Archives Initiative* (OAI), que definiu, entre outros aspetos, um código partilhado para *tags* de metadados e protocolos de interoperabilidade. Um dos resultados subsequentes à primeira reunião da OAI foi a adaptação de um *software* já existente, o CogPrints, para facilitar a criação de um novo tipo de repositório: os repositórios institucionais. Esta plataforma foi designada EPrints e foi apresentada publicamente em 2000. Desde então, foram surgindo outros sistemas e plataformas para a criação de repositórios como o DSpace, o Fedora, entre outros (SDUM, 2015).

A declaração de Budapeste marca o início do que movimento pelo acesso aberto à informação científica. Assim, a 14 de fevereiro de 2002, a *Budapest Open Access Initiative* declara a importância da remoção das barreiras de acesso à literatura científica através de duas estratégias complementares: o autoarquivo em repositório – a via verde – , ou a publicação em uma revista de acesso aberto – a via doutorada. De fato,

> Desfazer as barreiras que impedem o acesso a esta literatura irá acelerar a pesquisa, fortalecer a educação e difundir o conhecimento de maneira geral, tirando dela seu máximo proveito e assentando as bases para a união da humanidade em uma ampla e inédita conversação intelectual comum em sua marcha pelo conhecimento. (BOAI, 2002).

Um RI é um tipo de biblioteca digital, predominantemente de natureza acadêmica, que coleta, dissemina e preserva a produção intelectual de uma instituição incentivando a participação da sua comunidade no movimento pelo acesso aberto (Burns, Lana e Budd, 2013; SDUM, 2015). Um RI, quando permite o acesso aberto ao seu conteúdo e ao *harvesting* dos seus metadados, é tipificado como um repositório de acesso aberto (Heery e Anderson, 2005).

A literatura de acesso aberto é digital e as cópias dos documentos científicos são disponibilizadas na internet, sem custo para o autor nem para a sociedade. Também são livres da maioria das restrições de *copyright* e licenças (Suber, 2015). O acesso aberto é útil e desejável, porque quebra o paradigma da comunicação científica das revistas pagas, uma vez que esta modalidade de publicação limita o acesso às informações científicas para uma significativa parcela de pesquisadores ao redor do mundo. Além disso, as restrições inerentes às publicações pagas causam uma perda de eficiência considerável no sistema de comunicação científica, uma vez que limitam o impacto e o reconhecimento dos resultados obtidos pelos pesquisadores e pelas instituições onde eles realizam suas atividades. Logo, a disponibilização da literatura científica em acesso aberto é imprescindível a um sistema de comunicação científica para que este impulsione o progresso da ciência e sua eficácia (SDUM, 2015). Atente-se, contudo, para o fato de que o acesso aberto não significa que não haja um custo de produção. Dentre as várias

explicações elencadas por Suber (2015), destacamos a assertiva de que a economia na produção da literatura de acesso aberto, comparada com as publicações convencionais, e o modelo de negócios, para recuperar os custos, dependem do lócus onde a literatura vai ser disponibilizada, ou seja, em revistas ou repositórios de acesso aberto. Essas duas plataformas que disponibilizam os documentos científicos, por sua vez, concretizam o acesso aberto.

1.3 − O repositório institucional *Open Access*: componentes e políticas

Um RI do tipo *Open Access* é um sistema que viabiliza a chamada via verde do movimento *Open Access*. O sucesso[6] do funcionamento de um sistema como esse depende de um processo de estímulo e conscientização da comunidade de produtores de objetos digitais a serem armazenados, com destaque especial para os autores de artigos científicos. No entanto, Borges (2006) adverte que antes de uma instituição se preocupar com as estratégias de adesão a esse novo meio de comunicação científica, torna-se prioritário estabelecer as condições de operacionalidade do sistema. Para tanto, a autora esquematizou um diagrama conceitual que serve como ponto de partida para uma definição pormenorizada dos componentes cultural, organizacional e técnico do sistema (figura 3). O desafio representado pelo componente cultural é uma preocupação central para a constituição de um RI, já que a adesão à via verde depende do interesse dos autores dos trabalhos científicos. Entretanto, uma possível solução para esse problema seria envolver a própria universidade, as faculdades, os departamentos e os pesquisadores. Assim, o RI funcionaria como um instrumento de *marketing* para as universidades e para o sistema da avaliação da produção e atividades desenvolvidas pelas faculdades, institutos e departamentos. A mesma funcionalidade se aplicaria às faculdades e departamentos. Para os pesquisadores cabe não só a vinculação automática de sua produção ao sistema de avaliação docente, como também o provável aumento do impacto de suas publicações.

[6] Um exemplo de sucesso na implementação de um RI foi relatado por Ferreira *et al.* (2008).

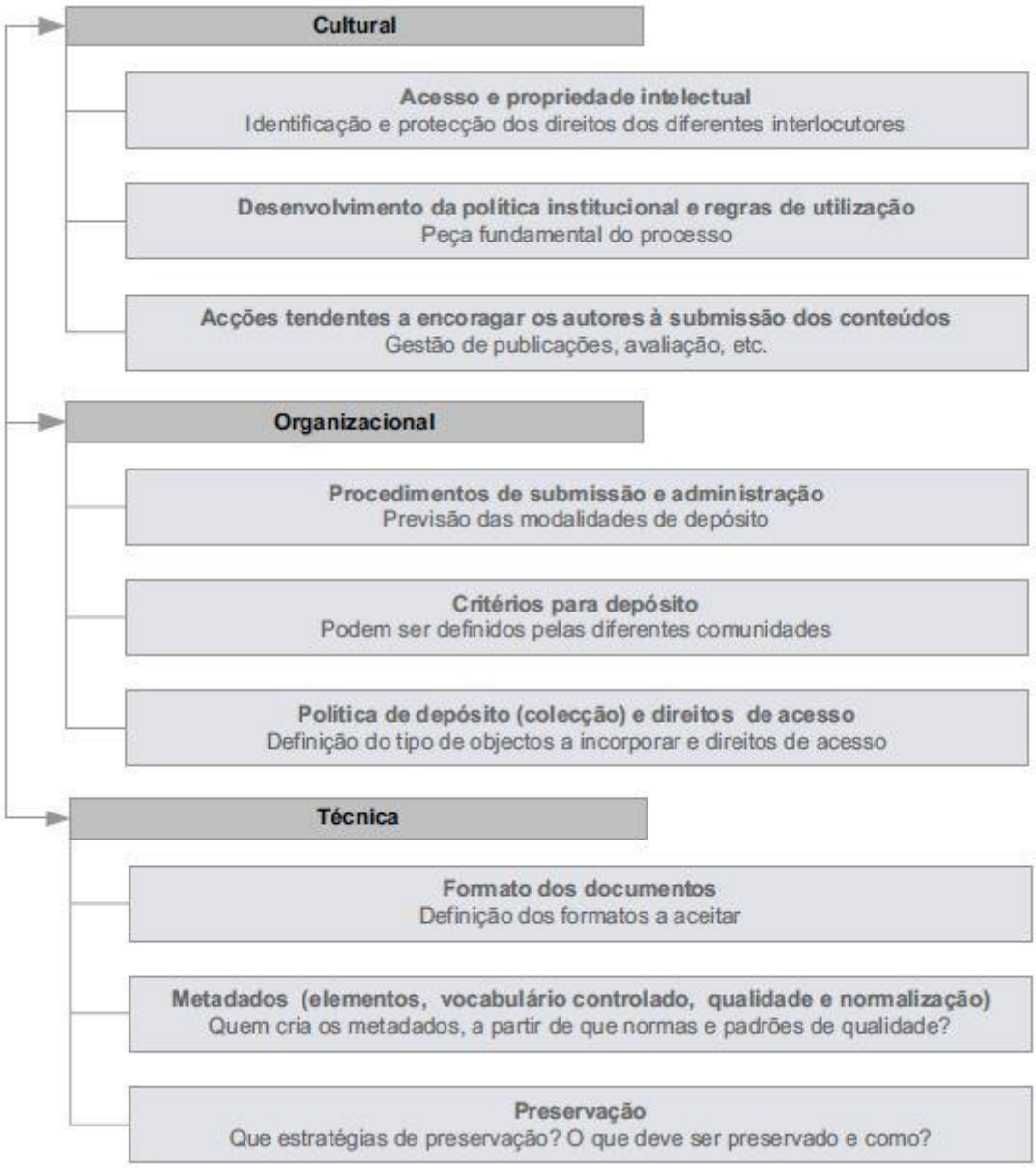

Figura 3 – Componentes de um repositório institucional

Fonte: Borges, 2006, p. 240

Podemos ver, no esquema da figura 3, que o desenvolvimento de uma política institucional e as regras de utilização são peças fundamentais no processo de constituição de um RI. De fato, uma política institucional é estabelecida por meio de portarias, resoluções e demais instrumentos legais, o que compromete a instituição e a comunidade universitária, com relação aos direitos e deveres de ambas as partes. Naturalmente, haverá políticas diferenciadas, como a que estabelece regras para o depósito e para os direitos

autorais, previstas no componente organizacional, além das que definem critérios de preservação digital, previstas no componente técnico.

Um RI tem o potencial para elevar o papel científico, social e acadêmico de uma universidade, pois:

> (...) Ao concentrar a produção intelectual dos investigadores de uma universidade, este torna-se o espelho dessa produção que representa o seu valor científico, social e financeiro. Um repositório institucional é mais do que um arquivo da produção intelectual de uma universidade, constitui uma afirmação de partilha de resultados científicos, a participação num esforço conjunto na constituição da ciberciência. (Borges, 2006, p. 464).

Ao corroborarmos essa assertiva, constatamos o fato de que várias universidades do mundo todo demonstram ter consciência do papel dos seus repositórios digitais na forma de institucionalização, por meio de instrumentos legais (políticas, portarias, resoluções etc.). Sem nos imiscuirmos na pertinência de suas respectivas políticas de informação, citamos dois exemplos: o Repositório Lume da Universidade Federal do Rio Grande do Sul (Brasil) e o RepositóriUM da Universidade do Minho (Portugal).

A política de depósito em um RI deve prever a aceitação do material produzido pelas diversas instâncias de uma universidade (faculdades e departamentos), requerendo dos depositantes uma simples afiliação. Em alguns casos, o material poderá ser submetido a uma revisão. Um RI tem como característica a acumulação de grande quantidade de objetos digitais, ano após ano. Isso requer um cuidadoso planejamento de preservação e acesso a longo prazo. Como mecanismo de acesso e recuperação de informação, um RI deve prover um sistema de busca próprio e metadados para viabilizar o serviço de *harvesting*. Uma política de acesso ao conteúdo irá estabelecer as regras adequadas às diferentes práticas disciplinares, com consideráveis vantagens para os pesquisadores e docentes em suas atividades didáticas. (Borges, 2006). Neste aspecto, um repositório se diferencia dos ambientes virtuais de ensino, porque estes ambientes têm acesso relativamente restrito. Eles funcionam como um acervo de materiais de ensino-aprendizagem, mas não possuem funções de busca e recuperação requeridas por um repositório. Tornando-se o material didático mais acessível, ainda que tal acesso seja restrito a uma determinada instituição, procede-se ao desafio para as instituições disponibilizarem sua produção intelectual publicamente (JISC infoNet, 2012).

O Diretório de Repositórios de Acesso Aberto (OpenDOAR, 2014) alerta para a necessidade de uma categorização dos repositórios com informações esclarecedoras acerca de suas respectivas políticas sobre os materiais revistos, ou não, pelos pares, assuntos cobertos, público-alvo, políticas de coleta e preservação etc. Onde esse tipo de informação não existir, os repositórios deverão ser encorajados a providenciá-lo para aumentar sua visibilidade e o uso dos objetos neles armazenados. O projeto OpenDOAR tem como principal financiador o *Joint Information Systems Committee* (JISC), que inclui em seus projetos as questões relacionadas com os repositórios de acesso aberto no Reino Unido. O JISC também financiou o *Repositories Supported Project* (RSP). Este projeto ajudou a construir competências e conhecimentos sobre repositórios nas instituições de ensino superior do Reino Unido. Apesar de o projeto ter sido encerrado em agosto de 2013, o seu *website* é uma excelente fonte para orientar as instituições de ensino nas linhas mestras da criação de repositórios de acesso aberto. Além disso, estão disponíveis dezenas de publicações relacionadas com a problemática dos repositórios. A seção *Policies & legal issues* (RSP, 2013), por exemplo, trata dos seguintes temas:

a) Políticas de conteúdo: sugestões para a criação de políticas que definirão o tipo de conteúdo que será armazenado no repositório.
b) Políticas de submissão: definição de políticas para receber conteúdos.
c) Políticas de reutilização de dados: sugestões de como o conteúdo de um repositório pode ser utilizado por terceiros.
d) Políticas de preservação: ajuda na formulação de uma abordagem sobre preservação para repositórios.
e) Questões de direitos autorais (*copyright*): segundo o RSP, as questões relacionadas ao *copyright* são as mais difíceis de resolver. Essa seção orienta a criação de políticas de *copyright*.
f) Políticas de remoção: o RSP alerta para a necessidade de uma política robusta para resolver conflitos sobre itens que foram submetidos. Essa seção orienta a definição de políticas de remoção de material disponível em um repositório.
g) Embargos: orientação na criação de políticas de conteúdo onde um embargo tenha sido interposto.

Um RI de acesso aberto tem importância equivalente às publicações de acesso aberto nas condições estabelecidas pelas agências de fomento à pesquisa, porque estas

requerem que as publicações dos resultados das pesquisas financiadas sejam publicadas em periódicos de acesso aberto ou em repositórios de acesso aberto. No tocante a esta última forma de publicação, isto é, a forma de arquivamento em acesso aberto (*Open Access Archiving*), uma política de depósito ideal possui três critérios básicos (SHERPA-JULIET, 2015):

a) As publicações devem ser depositadas em um repositório de acesso irrestrito e sem qualquer custo.
b) A versão publicada ou a versão final, *peer-reviewed*, deve ser depositada.
c) A concordância com os termos de uso deverá ocorrer no momento de aceitação para a publicação.

O projeto *Open Access Infrastructure for Research in Europe* (OpenAIRE, 2015) sugere vários documentos que tratam de recomendações para criação de uma política *Open Access*. Dentre eles, destacamos: o *Guide to good practices for university open-access policies* (Shieber e Suber, 2004) produzido pelo projeto *Open Access* liderado pela Universidade de Harvard, com participação de dezenas de universidades norte-americanas; e o *Open Access Policy Kit* (Rodrigues, 2009) produzido pelo projeto RCAAP e recomendado pelo *Harvard Open Access Project*.

No Brasil, o IBICT articulou um projeto de lei para obrigar as universidades e institutos de pesquisa públicos a criarem um RI. O projeto tem duas propostas básicas:

> (...) 1) estabelece a obrigatoriedade de as universidades e institutos de pesquisa públicos a construírem os seus respectivos repositórios institucionais e ao mesmo tempo torna obrigatório que os pesquisadores dessas universidades depositem uma cópia dos resultados de suas pesquisas, publicados em revistas com revisão por pares, nos referidos repositórios institucionais; 2) estabelece no seu artigo segundo que seja criada uma comissão de alto nível para a discussão e estabelecimento de política nacional de acesso livre à informação científica (Kuramoto, 2009, p. 213).

Em 25 de fevereiro de 2015, a situação da tramitação do projeto estava pronta para pauta na Comissão de Constituição e Justiça do Senado Federal, com voto pela rejeição por injuridicidade e inconstitucionalidade, no entendimento do relator do processo (Rollemberg, 2011). Contudo, dezenas de universidades públicas criaram seus RIs ao longo dos últimos anos, embora nem sempre precedidos por uma política de informação de acesso aberto ou como o desenrolar de uma política estratégica de tecnologia da

informação. Um exemplo de RI criado por uma iniciativa voluntária e individual é o da UFPB.[7]

[7] Relatos de experiência de implementação de RIs em universidades públicas do Brasil são descritos em Sayão *et al.* (2009).

Capítulo 2 − Aspectos conceituais da preservação digital

Neste capítulo, iremos estudar o catálogo de elementos de política de preservação do projeto SCAPE (Sierman, Jones e ElstrØm, 2014), com o objetivo de expandirmos os aspectos conceituais designados por tais elementos de preservação digital.

2.1 − Autenticidade

Segundo Duranti (1989), a origem da diplomática está diretamente relacionada com a necessidade de atestar a autenticidade de um documento. Todavia, a autenticidade diplomática se diferencia da autenticidade legal, mesmo que elas possam constituir uma atribuição de autenticidade histórica em uma disputa judicial. Esses três tipos de autenticidade se diferenciam da seguinte forma:

- Um documento legalmente autêntico é chancelado por uma autoridade pública garantindo sua genuinidade;
- Um documento diplomaticamente autêntico é aquele que foi escrito de acordo com a prática do tempo e lugar indicado no texto e assinado pelo autor;
- Um documento historicamente autêntico é aquele que atesta eventos que aconteceram ou informações verdadeiras.

A definição arquivística de autenticidade não se preocupa com a verdade do conteúdo de um documento. Não obstante, os processos arquivísticos pretendem estabelecer a autenticidade de modo que fortaleça a confiança entre o usuário dos objetos arquivados e a instituição responsável pela preservação desses objetos. Em um plano ideal os arquivistas mantêm os documentos diplomaticamente autênticos e os tornam acessíveis. Assim, as pessoas interessadas em registros culturais poderão utilizar estes documentos e determinar a verdade histórica dos seus respectivos conteúdos. Por outro lado, essa definição está estreitamente relacionada com o conceito de integridade do objeto. Se um objeto arquivístico possui integridade, subentende-se que ele seja completo e inalterado. A integridade diz respeito à comparação do objeto com a sua forma original, enquanto a autenticidade trata de verificar se o objeto é ou não o que se afirma sobre ele

durante o processo de seleção e avaliação. Assim, a informação de proveniência associada ao objeto é utilizada como um ponto de verificação, mas o conteúdo e os atributos físicos também podem ser considerados. Por exemplo: uma assinatura pode servir como um teste de autenticidade, identificando-se o criador e estabelecendo-se o seu relacionamento com o registro. Uma vez estabelecida a autenticidade de um objeto, o arquivista responsável pela preservação do objeto deverá garantir-lhe a integridade contra as ameaças ambientais e de segurança, assim como a autenticidade inalterável (Adam, 2010).

Ao analisar as medidas de integridade e a autenticidade no mundo analógico e digital, Seadle (2012) assegura que no ambiente digital não há originais; apenas cópias. Além disso, a caracterização de autenticidade torna-se um desafio devido à mutabilidade dos objetos digitais. Uma página *web*, que é a cópia de um código enviado em pacotes do servidor para os computadores dos internautas, pode ser apresentada diferentemente, quando é exibida em navegadores, computadores e telas diferentes. Essa mutabilidade implica dizer que, do ponto de vista da ciência da computação, a autenticidade está no código.

A autenticidade está muito mais relacionada com a integridade no espaço digital do que no ambiente analógico. Assim, um trabalho digital existente de modo inalterado em várias cópias possui uma forma de integridade que lhe garante a autenticidade, mas um objeto digital cuja integridade tenha sofrido alguma mudança terá sua autenticidade comprometida. Uma outra medida de autenticidade é a proveniência. Entretanto, quando esta se torna problemática, suscita questões semelhantes às do ambiente analógico. Mesmo assim, um teste de integridade pode dissipar as dúvidas, desde que uma cópia genuína possa ser utilizada para comparação.

No meio digital, a autenticidade pode significar uma cópia exata de um texto ou de uma imagem, mas também pode ser o texto e a imagem no seu contexto original. Isto pode ser exemplificado com o conflito de *copyright* ocorrido pela utilização de tiras cômicas de Dilbert no *website* de um professor universitário[8]: As cópias estavam com ligação para as imagens originais no *site* da empresa *United Media*. Contudo, a empresa considerou que houve uma alteração ilegal do contexto das imagens e, assim, reproduzia-se uma cópia inautêntica.

[8] *Dilbert Hack Page Archives*. Disponível na WWW: http://www.cs.rice.edu/~dwallach/dilbert/

Enquanto houver dificuldades para se estabelecer, de forma clara, uma medida de autenticidade, Seadle (2012) considera razoável determinar a autenticidade de um objeto digital quando a sua integridade puder ser mensurada e demonstrada como a mesma de um outro objeto num servidor seguro, por exemplo. Esse foi o caso do conflito de propriedade sobre as tiras de Dilbert. A integridade também atenderia aos fins de prova para atestar a proveniência e atestaria a autenticidade do conteúdo dos objetos digitais.

Doutra parte, o conceito de integridade de um objeto digital é um pouco mais complicado do que o de um objeto analógico. Se uma página de um livro de biblioteca for subtraída, compromete-se a integridade da obra, mas isso só será percebido quando alguém acusar a falta da página. De todo modo, o setor de restauração poderá reinserir uma cópia da página sem muita dificuldade, se não for uma obra rara. De qualquer forma, a integridade digital precisa de uma medida apropriada. Geralmente, os sistemas computacionais utilizam *checksums* ou *hashes* para verificar se dois objetos digitais são idênticos. Entrementes, usar essas técnicas como medida de autenticidade é problemático, porque os algoritmos subjacentes se caracterizam pela seguinte alternância: um arquivo tem integridade de acordo com a medida verificada ou não tem integridade. Isso significa que a integridade no ambiente digital é rígida, isto é, não possui a flexibilidade para interpretá-la no ambiente analógico.

Um outro modo de testar a autenticidade é executar um programa-padrão que consiga abrir um arquivo digital. Um processador de texto consegue ler um arquivo de caracteres ASCII puro. Ele só vai acusar algum erro se houver caracteres que não forem desse tipo. Qualquer alteração que for codificada em ASCII não será detectada, o que não ocorreria se fosse utilizado um *checksum*. Pequenos danos também podem passar despercebidos se não houver um escaneamento de todo o arquivo na tela.

Seadle (2012) finaliza seu estudo afirmando que um sistema poderia isolar uma área alterada em um arquivo e comparar outras partes para testar a sua integridade. Um sistema inteligente poderia também julgar o significado de uma alteração para qualificar uma medida de integridade. Para o autor, muitas questões-chave envolvendo a autenticidade digital ainda não foram testadas, mas isso ocorrerá quando o valor econômico de um trabalho digital autêntico tiver de ser discutido nos tribunais.

Os conflitos jurídicos não são totalmente evitáveis, mas uma instituição poderá preparar-se para este tipo de eventualidade se planejar a forma como pretende assegurar a autenticidade do material digital armazenado em seus repositórios. Um primeiro passo neste sentido seria o estudo de alguns documentos que fundamentam um *framework* básico para autenticidade e proveniência desenvolvido por Salza *et al.* (2012). Por exemplo:

- O projeto InterPARES, segundo esses autores, elaborou o principal *framework* conceitual relacionado com a autenticidade. Ele propicia a comparação e a avaliação da qualidade e consistência das práticas digitais, no que diz respeito à autenticidade.
- O modelo de referência OAIS (CCSDS, 2012) é um arquivo organizacional de pessoas e sistemas que aceitam a responsabilidade de preservar a informação e de garantir a uma comunidade alvo o acesso ao acervo.
- A norma ISO 16363 (2012) estabelece os critérios para auditar-se e certificar-se um repositório digital como confiável.
- A norma ISO 16919 (2014) estabelece os requisitos para as entidades que auditam e certificam os repositórios digitais.

Uma vez que o *staff* dos repositórios adquira conhecimentos como os enumerados acima, torna-se necessário elaborar uma política de orientação para se estabelecer a autenticidade dos objetos digitais. Essa política requer três elementos fundamentais: a integridade, a confiabilidade e a proveniência.

A integridade é verificada com medidas de integridade, tais como: encriptação, assinaturas digitais, verificação da fixidez etc. Uma política de preservação deve declarar que essas medidas irão contribuir para a autenticidade do material digital, tanto na ingestão como no armazenamento, dois aspectos relevantes explicados da seguinte forma:

> Ingest: The completeness of the digital object will need to be defined before ingest and could be part of the discussion with the content deliverer or producer. At ingest the received checksums can be compared with the checksums generated upon retrieval. This will show whether bits were lost during transportation. This measure should be implemented for all data movements, including when the data is moved inside the repository.
>
> Storage: moving data from one place to another needs to be accompanied by measures to check before and after the move whether the digital object is still complete and undamaged. This also applies to back up copies. Authentication

> measures to safeguard that personnel cannot make changes to the data stored or (unintentionally) delete (part of) digital objects (Sierman *et al*., 2014, p. 20).

A confiabilidade[9] é a "(...) credibilidade de um documento arquivístico enquanto uma afirmação do fato. Existe quando um documento arquivístico pode sustentar o fato ao qual se refere, e é estabelecida pelo exame da completeza, da forma do documento e do grau de controle exercido no processo de sua produção." (CTDE, 2014, p. 13). Ao estabelecer esse aspecto na manutenção de registros e nos processos de arquivamento, uma instituição irá convencer a comunidade de um repositório que os objetos preservados são confiáveis.

A proveniência significa a origem, história da propriedade ou localização de um objeto. O termo é usado, principalmente, por profissionais que trabalham com coleções de obras de arte, gestão arquivística, biblioteconomia, computação e direito. Ele tem o propósito de confirmar ou reunir evidências do tempo, do lugar e, quando for necessário, da pessoa responsável pela criação, produção ou descoberta do objeto (Archives New Zealand, 2015). Quando se trata de um objeto digital, o dicionário PREMIS traz a seguinte definição para o termo proveniência:

> Documentation of processes in a Digital Object's life cycle. Digital Provenance typically describes Agents responsible for the custody and stewardship of Digital Objects, key Events that occur over the course of the Digital Object's life cycle, and other information associated with the Digital Object's creation, management, and preservation." (Premis Editorial Committee, 2015, p. 269)

O conhecimento das necessidades de uma comunidade de usuários facilita a escolha dos elementos de proveniência para estabelecer a autenticidade do material preservado. Um conjunto de dados que não contiver as informações sobre o *software* e os parâmetros utilizados para gerar tais dados comprometerá a autenticidade do ponto de vista de um pesquisador. Da mesma forma, a falta de informações sobre a editora de um *e-book* pode tornar o objeto digital inútil para um pesquisador literário. A comunidade-alvo também precisa examinar os elementos essenciais para validar a proveniência dos objetos digitais de suas coleções (Sierman *et al.*, 2014). A proveniência desempenha um

[9] Essa definição é igual à encontrada no DPGlossary que, por sua vez, a compilou do projeto InterPARES. Diponível na WWW:
http://www.alliancepermanentaccess.org/index.php/consultancy/dpglossary/#Reliability

papel importante nas ações de preservação, como a migração e a padronização, temas que serão discutidos nas próximas seções.

2.2 − Preservação de *bit*

A preservação de *bit* é definida pelo dicionário PREMIS nos seguintes termos:

> Preservation strategy in which the sole objective is to ensure that a Digital Object remains fixed (unaltered) and viable (readable from media). No effort is made to ensure that the Digital Object remains renderable or interpretable by contemporary technology (Premis Editorial Committee, 2015).

Para que um objeto digital permaneça inalterado, é preciso garantir sua fixidez[10] (*fixity*), ou seja, que ele não sofra modificações em um dado intervalo de tempo. Em outras palavras: a preservação de *bit* assegura que a cadeia de *bits* de um objeto digital permanece intacta e recuperável ao longo do tempo (Lavoie e Dempsey, 2004). Entretanto, a preservação de *bit* é apenas um subconjunto de aspectos da preservação digital. Na verdade, ela é o ponto de partida de todas as atividades da preservação digital. Por essa razão, torna-se necessário conceituá-la de modo que se incluam diversos aspectos da preservação digital por meio de uma abordagem holística da preservação de *bit*, que é definida como "(...) uma abordagem onde a preservação de *bit* é vista como algo que deve ser reconhecido como parte de um todo, onde diferentes circunstâncias podem influenciar em como a preservação de *bit* deve ser realizada" (Zierau, 2011, p. 11). Neste caso, o "todo" são todos os aspectos da preservação digital, conforme está ilustrado na figura 4.

[10] Para entender o conceito de fixidez (*fixity*), consultar o relatório *Checking your digital content: what is fixity, and when should I be checking it?* Disponível na WWW: http://www.digitalpreservation.gov/ndsa/working_groups/documents/NDSA-Fixity-Guidance-Report-final100214.pdf

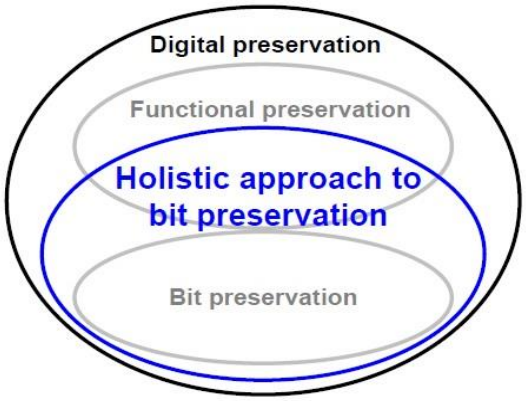

Figura 4 – Abordagem holística para a preservação de *bit* na preservação digital

Fonte: Zierau (2011, p. 11)

A abordagem holística para a preservação de *bit* abrange todos os aspectos da preservação de *bit* e apenas uma parte dos aspectos da preservação funcional. Esses dois conjuntos de aspectos influenciam na escolha das vias para assegurar a preservação de *bit*. Observa-se, ainda, na figura 4 que essa abordagem também inclui outros aspectos da preservação digital que não pertencem à preservação de *bit* ou à preservação funcional. Por exemplo: os aspectos de sustentabilidade. A abordagem holística não é um outro modo de descrever a preservação digital, mas um modo de incluir aspectos que influenciam na preservação de *bit*.

Em um dado momento, a preservação de *bit* deverá ser assegurada por uma solução de preservação de *bit*. A abordagem holística se concentra nos requisitos para uma tal solução, quando os *bits* do material preservado e a solução encontrada fazem parte de um todo. Esses requisitos podem equivaler aos requisitos para um subsistema organizacional escolhido como uma solução de preservação de *bit*. O subsistema será estruturado como um repositório de *bit*. Existem vários exemplos advindos do "todo" a demonstrarem que a preservação de *bit* pela abordagem holística não está restrita à integridade e à legibilidade dos *bits*, tais como: os aspectos da segurança da informação − integridade, disponibilidade e confidencialidade, conforme a norma ISO 27000; os requisitos de custo; os aspectos de representação do material digital preservado ao nível do *bit*, dentre outros.

A abordagem holística para a preservação de *bit* visa a contribuir para se otimizar uma forma de preservar o *bit*, com base em três resultados da pesquisa de Zierau[11] (2011): um modelo para definir a preservação de *bit*, separadamente de outros aspectos da preservação digital; um conceito para representações de material digital; uma metodologia para se avaliar a escolha de diferenciadas soluções de preservação de *bit*.

O tipo e a importância do material digital para uma dada coleção são determinantes para a escolha das ações de preservação. Assim, pode-se adotar um modelo de níveis de preservação digital, como o do quadro 3 adiante, para se escalonar a estratégia de preservação de *bit* associando-a com outros elementos essenciais do "todo".

Quadro 3 – Níveis de preservação digital

	Level 1 (Protect your data)	Level 2 (Know your data)	Level 3 (Monitor your data)	Level 4 (Repair your data)
Storage and Geographic Location	- Two complete copies that are not collocated. - For data on heterogeneous media (optical discs, hard drives, etc.) get the content off the medium and into your storage system.	- At least free complete copies. - At least one copy in a different geographic location. - Document your storage system(s) and storage media and what you need to use them.	- At least one copy in a geographic location with a different disaster threat - Obsolescence monitoring process for your storage system(s) and media	- At least three copies in geographic locations with different disaster threats - Have a comprehensive plan in place that will keep files and metadata on currently accessible media or systems.
File Fixity and Data Integrity	- Check file fixity on ingest if it has been provided with the content. - Create fixity info if it wasn't provided with content	- Check fixity in all ingests. - Use write-blockers when working with original media. - Virus-check high risk content.	- Check fixity of content at fixed intervals - Maintain logs of fixity info; supply audit on demand - Ability to detect corrupt data - Virus-check all content	- Check fixity of all content in response to specific events or activities - Ability to replace/repair corrupted data - Ensure no one person has write access to all copies
Information Security	- Identify who has read, write, move and delete authorization to individual files	- Document access restrictions for content.	- Maintain logs of who performed what actions on files, including	Perform audit of logs

[11] A discussão dos resultados da pequisa de Zierau (2011), entretanto, está fora do escopo de nossa tese.

	- Restrict who has those authorizations to individual files		deletions and preservation actions.	
Metadata	- Inventory of content and its storage location. - Ensure backup and non-collocation of inventory	- Store administrative metadata. - Store transformative metadata and log events	- Store standard technical and descriptive metadata	- Store standard preservation metadata
File Formats	- When you can give input into the creation of digital files encourage use of a limited set of known open formats and codecs.	- Inventory of file formats in use.	- Monitor file format obsolescence issues	- Perform format migrations, emulation and similar activities as needed

Fonte: Phillips *et al.*, 2013

Trata-se de um modelo de preservação digital que exclui as questões relacionadas com direitos e/ou com as políticas e se restringe às questões técnicas. Ele foi criado pela NDSA com o objetivo de ajudar os interessados no acesso à informação digital a longo prazo a avaliarem suas práticas para minimizar os riscos de perda, ajudando-os também a identificarem os próximos passos que deverão ser dados a fim de movimentar todas as operações, ou parte delas, para um próximo nível. Estas operações constituem as atividades desenvolvidas ao longo das fileiras de níveis. Quando ocorre a movimentação das fileiras, do nível 1 ao nível 4, tais operações deslocam-se da necessidade básica da preservação de *bit* em direção a requisitos mais abrangentes a fim de se rastrear o conteúdo digital e de tornar possível sua acessibilidade por um longo período de tempo.

O quadro 3 dispõe as linhas de orientação em cinco áreas funcionais ou categorias que compõem o núcleo dos sistemas de preservação digital:

a) Armazenamento e localização geográfica – O foco dessa categoria é o armazenamento da informação digital. A cada mudança de nível acrescentam-se novas cópias. Dessa forma, restringem-se as ameaças decorrentes de uma degradação de *bit* ou de falhas na mídia ou do sistema de armazenamento. A diversificação da localização geográfica em cada nível ajuda, por sua vez, na proteção contra ameaças ao sistema de armazenamento, seja por causa dos desastres naturais ou por aqueles provocados por seres humanos. Nos níveis 2, 3 e 4, para reforçarem a longevidade dos sistemas de armazenamento, foram

adicionados os seguintes requisitos: documentação do sistema, monitoramento da obsolescência e plano de acessibilidade. Assim, simplifica-se o trabalho, porque as atividades são desempenhadas a passos incrementais.

b) Fixidez de arquivo e integridade dos dados – O objetivo dessa categoria é o de prover uma série de passos que conduzirá uma organização a um estágio onde ela estiver atuando incisivamente para assegurar a fixidez (*fixity*) do seus objetos digitais. No nível 1, é requerido um *fixity check*[12] na ingestão. Caso não haja informação de fixidez, gerar-se-á uma. Esse processo é fundamental para se garantir que o conteúdo preservado é o que realmente se pretendeu preservar. Os níveis seguintes levam a uma operação contínua de *fixity check*, aumentando, assim, a confiança na fixidez do conteúdo que está aos cuidados preservacionistas de uma organização.

c) Segurança da informação – Esta categoria visa a identificar quem lê, escreve, executa e exclui um objeto digital, registrar *logs* de manipulação do objeto e atribuir restrições de acesso. Neste caso, pode-se prevenir, por exemplo, a exclusão acidental de um objeto digital. Os níveis são organizados por incrementos de ações que atendem a requisitos cada vez mais avançados e regulados para minimizar os riscos que também estão relacionados com os riscos percebidos nas demais categorias.

d) Metadados − Nesta categoria, são apresentadas as camadas adicionais de metadados, as quais tornarão o conteúdo mais protegido, identificável e acessível. Na maioria dos sistemas, quase todos esses metadados (exceto os descritivos) deveriam ser gerados e processados exclusivamente por computador.

e) Formato de arquivos − Os objetos digitais estão sujeitos à estrutura e tipo dos seus formatos de arquivo. O nível 1 sugere que as organizações procurem utilizar formatos de arquivos abertos e conhecidos; o nível 2 documenta os formatos em uso; o nível 3 monitora a obsolescência dos formatos; o nível 4

[12] Fixity check: a mechanism to verify that a digital object has not been altered in an undocumented manner; Checksums: message digests and digital signatures are examples of tools to run fixity checks; Fixity information: the information created by these fixity checks, provides evidence for the integrity and authenticity of the digital objects and are essential to enabling trust. Disponível na WWW: http://ndsa.org/glossary/

> recorre à migração, emulação de suporte e examina outros modos de assegurar que o conteúdo preservado estará usável e acessível no futuro.

Cabe observar que os formatos de arquivo são suscetíveis a erros de *bit*. Diferenciados formatos de arquivos irão requerer diferenciadas soluções de preservação de *bit*, de modo que alcançarão a mesma maximização de prevenção de riscos contra a perda de material digital[13] (Zierau, 2011). O dicionário PREMIS recomenda que o formato de arquivo seja apurado pelo repositório no processo de ingestão, uma vez que muitas atividades de preservação dependem do conhecimento detalhado sobre o formato do objeto digital. O *fixity check* e o *virus check* também devem ser realizados na ingestão, conforme as orientações da categoria fixidez de arquivo e dados íntegros no modelo de níveis de preservação digital. O modelo apresentado está na primeira versão, mas serve para orientar tanto aqueles que estão iniciando as ações de preservação dos seus ativos digitais quanto as instituições que planejam aperfeiçoar seus sistemas e *workflows* de preservação digital. Ele permite que as instituições avaliem o nível de preservação alcançado para determinado material sob sua custódia; todavia, o modelo não foi planejado para avaliar os programas de preservação digital em todos os seus aspectos, pois não aborda questões como as relacionadas com política de preservação, pessoal ou apoio institucional (Phillips *et al.*, 2013).

Um objeto digital, uma vez que esteja disponível na *Web*, deverá ser identificável e acessível a longo prazo, independentemente do seu *Uniform Resource Locator* (URL). Um URL tem o propósito de identificar um recurso e descrever sua localização, mas poderá tornar-se inconsistente se o recurso for movido para uma outra localização. Por essa razão, o uso de identificadores persistentes é considerado a melhor solução para preservar o acesso ao recurso digital, independentemente do seu URL, pois o identificador persistente será associado a uma nova localização quando o recurso for movido (Bellini *et al.*, 2012). Entretanto, existem poucas estratégias para a implementação de identificadores persistentes e elas dependem das condições técnicas, administrativas e políticas das instituições, suas visões de futuro e interesse na interoperabilidade com outros sistemas. Nomeadamente, as estratégias são as seguintes:

[13] Material digital é um conjunto de informações ou objetos digitais (Ferreira, 2006).

> a) Redirecionamento - é uma estratégia mínima, posto que utiliza os recursos padronizados do servidor web para redirecionar as solicitações para a posição corrente do recurso. Este método é difícil de se gerenciar quando se trata de websites de grande porte;
>
> b) Instalação de um resolvedor apoiado em banco de dados - pressupõe um software servidor de links, rodando sobre um banco de dados e tendo como finalidade, mapear a localização corrente do recurso, ou seja, o URL corrente. Uma opção nesta categoria é o software servidor PURL - Persistent URL - disponibilizado pela Online Computer Library Center (OCLC) (http://purl.oclc.org/) – (...);
>
> c) Contratação de sistema de identificação persistente, oferecido por outra organização - existem vários sistemas de identificação persistente projetados para uso na Internet, baseados em padrões abertos, com objetivos e enfoques distintos. Por exemplo: Digital Object Identifiers (DOI), Handle System e também PURL, posto que a OCLC oferece serviço de identificação on-line para terceiros. (...) (Sayão, 2007, p. 68).

Acrescente-se aos sistemas exemplificados acima o *Uniform Resource Name* e o *Archival Resource Key*. A identificação persistente criada para autores também possui várias opções de implementação, tais como: o *AuthorClaim*, o *Scopus Author ID*, o *Researcher ID*, o *arXiv Author ID* e o *ORCID*. O grupo de trabalho *Institutional Identifier* da *National Information Standards Organization* iniciou o desenvolvimento de um padrão de identificação persistente para as instituições. Por outro lado, apesar de essas iniciativas demonstrarem uma crescente conscientização e interesse pelos identificadores persistentes, algumas dificuldades continuam a fazer da identificação persistente um problema complexo, pois as diferenciadas comunidades de usuários não conseguem garantir a persistência dos seus sistemas identificadores, especificamente os resolvedores. Por exemplo: bibliotecários, arquivistas, pesquisadores, editores e agências financiadoras possuem diferenciadas visões e abordagens para conceituar Identificadores Persistentes, bem como diferenciados modelos de negócio, critérios legais, requisitos e políticas.

Consequentemente, alguns sistemas identificadores acabam por abordar melhor as necessidades de determinadas comunidades, mas muitas dessas soluções particulares são largamente utilizadas para atender a requisitos específicos. Isso significa que a discussão sobre Identificadores Persistentes não pode restringir-se aos aspectos técnicos para assegurar a identificação persistente aos recursos digitais, pois trata-se de se levar em conta a complexidade de uma gama de responsabilidades e requisitos, os quais fundamentam o desenvolvimento e a manutenção de um sistema identificador. Cada um desses requisitos envolve o comprometimento de muitos *stakeholders* para manter uma

infraestrutura adequada e garantir um consenso em políticas, responsabilidades, direitos e deveres.

A dificuldade em se estabelecer a interoperabilidade desses sistemas pode residir em questões de ordem financeira. Como a criação de um único identificador global está longe de ser adotada, resta o desafio de estabelecer um *framework* de interoperabilidade entre as soluções de identificadores persistentes conhecidas para habilitar o acesso persistente, reuso e troca de informação por meio do uso dos identificadores existentes e recursos associados através de sistemas diferentes, localizações e serviços. Dessa forma, ninguém ficaria dependente de um único sistema identificador (Bellini *et al.*, 2012).

Sierman *et al*. (2014) recomendam que uma instituição defina o que ela entende sobre a preservação de *bit* e a relação com a preservação funcional. Essa definição, declarada em uma política de preservação digital, poderia incluir as ações descritas no quadro 3 de acordo com os níveis de preservação escolhidos para determinados tipos de material digital. A fim de assegurar a performance da preservação de *bit*, uma instituição deve tomar algumas precauções no processo de ingestão desse material:

- Verificar a existência de vírus antes da ingestão.
- Certificar-se de que a coleção e os objetos estão completos.
- Identificar, caracterizar e validar os formatos.

Do mesmo modo, é importante que a política de preservação digital inclua a atribuição de identificadores persistentes ao material digital − pois eles minimizam o risco de perdas causadas pela falta de identificação − determine o número de cópias dos objetos digitais, a distribuição geográfica –, isto implica delegar a administração do material distribuído a outra equipe – e tenha um plano de contingência para a recuperação de desastres.

2.3 − Preservação funcional

A preservação funcional é a preservação de algumas ou todas as funções do ambiente do *software* original. A preservação de *bits* será inútil se não for possível decodificá-los e usar a informação. Por conseguinte, a preservação das funções da

aplicação original será o próximo nível de um programa de preservação. Mesmo assim, nem sempre se faz necessária a preservação de toda a funcionalidade da aplicação que gerou a informação (Waugh *et al.*, 2000).

A preservação funcional assegura a permanência da inteligibilidade e da utilidade dos *bits* de acordo com os propósitos da preservação (Zierau, 2011), os quais são condicionados pela política institucional (Smith *et al.*, 2003). O desafio desse tipo de preservação já se apresentava como preocupação em trabalhos de estudiosos, como Hedstrom (1998), Kuny (1998), Garrett eWaters (1996). Dentre os vários motivos que contribuem para a obsolescência da informação digital, destacam-se as rápidas mudanças nos meios de recodificação, a crescente variedade de formatos de arquivos e funcionalidades integradas de documentos e a execução em diferenciadas plataformas, conforme atestam os estudos de Gladney (2007), Rauch e Rauber (2004) e Thibodeau (2002). Entretanto, diversos outros aspectos da preservação funcional têm sido objeto de pesquisa ao longo das duas últimas décadas, entre os quais ressaltamos os relacionados com o planejamento da preservação:

> Preservation Planning, i.e. evaluating preservation strategies and choosing the most appropriate strategy, has turned into a crucial decision process, depending on both object characteristics as well as institutional requirements. The selection of the preservation strategy and tools is often the most difficult part in digital preservation endeavours; technical as well as process and financial aspects of a preservation strategy form the basis for the decision on which preservation strategy to adopt. The area of Preservation Planning has therefore attracted much interest in recent years (Strodl *et al.*, 2007, p. 4).

Um exemplo desse interesse é o trabalho de Becker *et al.* (2009), no qual os autores diferenciam os conceitos de política de preservação digital e o de um plano de preservação para, então, descrever os elementos necessários para formular um plano fundamentado e completo.

Existem diversas metodologias[14] para a preservação funcional, mas três delas são consideradas as principais: a migração, a emulação e a preservação da tecnologia.[15]

---

[14] Outras estratégias são abordadas no verbete *Digital Preservation* em: http://en.wikipedia.org/wiki/Digital_preservation. Este verbete é mantido pelo projeto *Digital Preservation on Wikipedia*, que, por sua vez, é coordenado pelo Grupo de Trabalho *Standards and Practices* da NDSA na *Library of Congress*

[15] Pearson e Pozo (2009, p. 4) desenvolveram uma abordagem prática das metodologias de preservação digital, estruturando-a de acordo com os seguintes tópicos: "What the methodology is, or what it purports

a) Migração – Esta metodologia consiste na transferência periódica dos materiais digitais de uma configuração de *hardware/software* para uma outra configuração, ou de uma geração de tecnologia de computador para uma outra geração. O propósito da migração é o de preservar a integridade dos objetos digitais e permitir que os usuários possa recuperá-lo, exibi-los e utilizá-los, mesmo diante das constantes transformações tecnológicas (Garrett e Waters, 1996). Como a cadeia de *bits* é alterada após a migração, é preciso preservar também a autenticidade do objeto para que a estratégia tenha êxito (Zierau, 2011). Na migração, um formato de arquivo original é recebido e, a seguir, é trocado por um outro formato. A estratégia de migração a ser utilizada vai depender de cada caso. Por exemplo: alguns tipos de arquivo, em uma dada coleção, podem ser mais adequados para uma migração na ingestão, enquanto outros podem ser mais apropriados para uma migração baseada nos aspectos de risco. Como demonstram os estudos de Pearson e Pozo (2009) e Lee *et al.* (2002), essa metodologia apresenta mais desvantagens do que vantagens, porém, é uma das mais utilizadas, conforme asseguram Zierau (2011) e Ferreira *et al.* (2012).
b) Emulação – Esta metodologia é o processo de criar uma versão virtual do ambiente original que era utilizado para acessar um determinado objeto. O ambiente virtualizado é acessado por meio de um emulador executável em uma moderna plataforma de *hardware* e *software*. Isso permite que o acesso ao conteúdo original seja mantido por meio do emulador, sem que o conteúdo sofra alterações. A emulação não mantém a mesma forma e performance do *hardware* original. Isso pode ter implicações que dependem do planejamento da preservação. Quando a emulação, por si só, não permite o acesso adequado ao conteúdo dos materiais digitais, se faz necessário associá-la a uma outra metodologia, como a dos *renderers*. Na emulação (figura 5), um ambiente de *software/hardware* original é encapsulado em um ambiente emulado (*emulated environment*). Esse novo ambiente é instalado em uma plataforma

to do; How it works; Its perceived advantages and disadvantages; and Different strategies for approaching or maintaining the methodology."

contemporânea de *hardware* e *software*. Uma vez instalado na nova plataforma, utiliza-se o ambiente emulado para acessar o objeto pretendido.

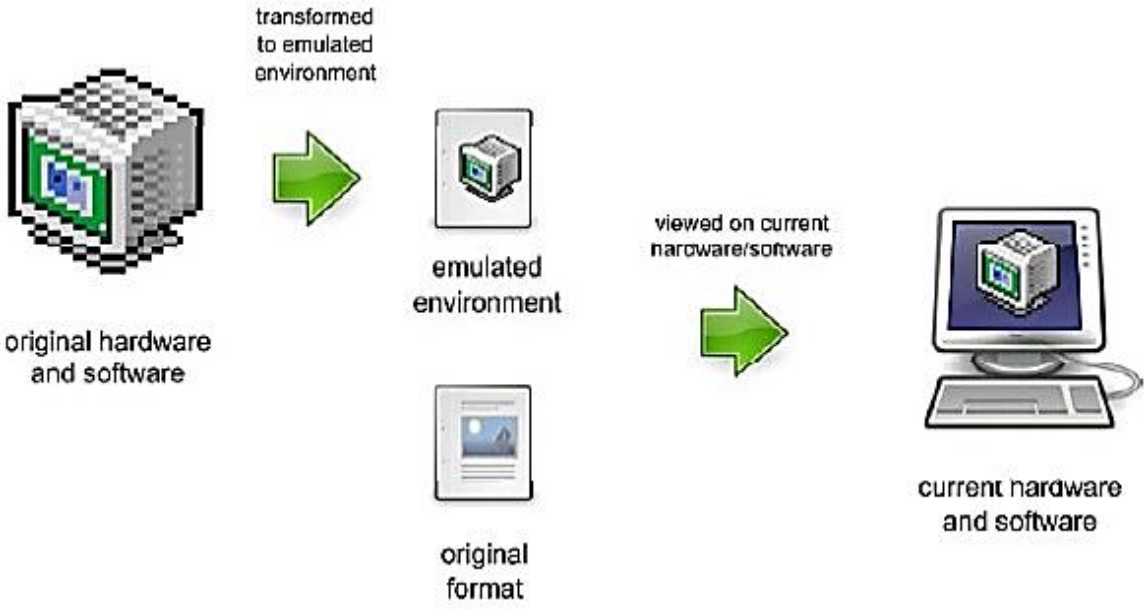

Figura 5 – Emulação
Fonte: Pearson e Pozo (2009, p. 22).

c) Preservação da tecnologia − esta estratégia consiste em preservar a plataforma de *software* e *hardware*, onde o material digital foi criado ou era acessado. Segundo Ferreira (2006), trata-se da criação de um museu de tecnologia, onde o alvo da preservação é o formato original do objeto digital. Os defensores dessa estratégia consideram-na a única forma de manter a fidedignidade dos objetos digitais. O *Handbook of Digital Preservation* (DPH, 2015), Beagrie e Jones (2008), Pearson e Pozo (2009) listam as vantagens e desvantagens dessa estratégia. Entre as vantagens, destacamos a de poder ser a única opção de se fazer a leitura de materiais digitais. Não obstante, as plataformas de *hardware* e *software* estão sujeitas à extinção provocada pela obsolescência tecnológica, o que caracteriza a desvantagem mais relevante. Conforme o *Handbook of Digital Preservation* (DPH, 2015), essa estratégia impõe os seguintes requisitos:

- Políticas e diretrizes relativas ao acesso.

- Documentação do *hardware* e *software* mantidos.
- Metadados requeridos para manter o *hardware* e o *software*.

2.4 − Objeto digital

Um objeto digital é um conjunto de uma sequência ou mais de *bits* que registram o conteúdo do objeto e o de seus metadados associados. Ele está estruturado em três níveis: o físico, o lógico e o conceitual (CTDE, 2011). No nível físico, o objeto está em um formato incompreensível para humanos. No nível lógico, tem-se o objeto produzido em uma linguagem de programação de alto nível que provê o código do objeto, ou seja, esse objeto se for um texto, pode ser do tipo *.rtf*, *.pdf*, *.doc* etc; se for um áudio, *.wav*, *.mp3*; se for uma imagem, *.jpeg*, *.png* e assim por diante. No nível conceitual, os objetos ganham forma quando se utilizam programas capazes de traduzir os vários tipos de formato de arquivos. Assim, poderemos ler um texto, visualizar uma imagem ou ouvir uma música.

No processo de ingestão do objeto digital, o repositório deverá garantir a integridade da informação recebida e produzir a informação descritiva, que irá possibilitar a localização e recuperação do objeto armazenado. Entretanto, esse objeto poderá ser recebido, em seu formato original, como um substituto do original ou como uma cópia. Assim, define-se o objeto original como o objeto recebido, isto é, *the primary authentic and unique item, either the original or the closest surviving surrogate or copy, as originally acquired by the Library.*[16]

Durante o ciclo de vida de um objeto digital, diversas ações de preservação podem ser adotadas para manter acessível e autêntico o objeto recebido. Uma delas poderia resultar em um novo objeto digital, como na migração de formatos de arquivos. Entrementes, nem sempre isso é viável, ainda que se mantenha o objeto original no repositório e as ações de preservação sejam feitas apenas nas cópias. Por essa razão, as instituições precisam desenvolver uma política que descreva como elas irão lidar com a

[16] Definição do termo *received version* no glossário da NDSA. Disponível na WWW:< http://ndsa.org/glossary/>

migração das cópias. Por exemplo: quando várias migrações fossem realizadas, apenas o objeto original e uma nova versão seriam preservados e nenhuma das versões intermediárias seria mantida.

Por outro lado, um objeto digital poderá vir a ser eliminado por diversas razões. Assim, uma política de preservação deverá prever o descarte ou a eliminação planejada dos objetos digitais. Essas ações seriam decorrentes de acordos de preservação ou efeitos de mandatos legais, como os que ocorrem no âmbito das instituições arquivísticas.[17] Seja qual for a regulamentação adotada, a comunidade de usuários deverá ser devidamente informada sobre as razões da remoção de um dado objeto ou coleção digital. Gerenciar descartes é uma atividade imprescindível na administração de coleções, porque existem razões específicas de um dado objeto (ou coleção) ter sido removido, tais como uma mudança na missão institucional, uma mudança de lugar da coleção ou, possivelmente, a instituição operacionalize uma política de retenção e descarte[18] (Sierman *et al.*, 2014).

Como foi visto na explicação do quadro 3, item 'e', os objetos digitais são moldados pela estrutura e tipo dos seus formatos de arquivo.[19] Assim, recomenda-se às instituições que prefiram os formatos de arquivos abertos e conhecidos,[20] que eles sejam documentados, monitorados com relação a obsolescência dos formatos e sofram ações de preservação por meio de estratégias de migração, emulação, dentre outras, a fim de se assegurar o uso e o acesso ao conteúdo a longo prazo.

No caso do monitoramento dos formatos de arquivo, uma instituição deve aplicar uma estratégia para manter as informações sobre os formatos e suas diferentes versões nos metadados dos objetos digitais e ao mesmo tempo definir como será feito o

[17] Um Sistema Informatizado de Gestão Arquivística de Documentos deve prever as seguintes ações: retenção dos documentos à prazo determinado, eliminação, transferência e recolhimento. Todas essas ações serão amparadas por normas e leis e estão definidas no documento e-ARQ Brasil (CTDE, 2011).

[18] Um exemplo de política de remoção de coleções digitais é a do UK Archive Data. Disponível na WWW:http://data-archive.ac.uk/media/54776/ukda062-dps-preservationpolicy.pdf .

[19] Um objeto digital pode ser classificado como simples ou complexo. Esta classificação está disponível na WWW: <http:// http://www.alliancepermanentaccess.org/index.php/consultancy/dpglossary/#Digital_Object>

[20] Um exemplo de determinação de uso do formato aberto está na política de preservação digital do National Archives of Australia (NLA, 2015). Disponível na WWW:< http://www.naa.gov.au/about-us/organisation/accountability/operations-and-preservation/digital-preservation-policy.aspx#section8>

acompanhamento contínuo do desenvolvimento dos formatos. Em alternativa, uma instituição poderá valer-se de um serviço externo para este fim, ou seja, de um registro central de formatos.[21] De todo modo, o monitoramento das versões dos formatos dos arquivos é imprescindível para possibilitar a melhor escolha e aproveitamento das estratégias de preservação funcional.

Em algumas circunstâncias, uma instituição poderá ter de remover um objeto digital ou uma coleção, seja porque esse material não estará mais disponível para a comunidade-alvo seja porque ele será retirado do repositório. Dessa forma, se faz necessária a criação de uma política de remoção[22] para deixar claro para a comunidade-alvo e produtores de material digital quais são as ações previstas com relação à existência futura dos objetos (Sierman *et al*, 2014).

A preservação dos objetos digitais precisa ser iniciada logo após a ingestão e prosseguir continuamente por meio de ações de intervenção. As intervenções realizadas por meio das estratégias de migração − chamada de *transformation*s no modelo de referência *Open Archival Information System* (OAIS) − resultam em alterações e possíveis perdas de dados, o que faz persistir a questão sobre o que se deve preservar (Grace, Knight e Montague, 2009). A resposta a esse tipo de questionamento se inicia pelo entendimento do que vem a ser as propriedades relevantes (*significant properties*) de um objeto digital. Assim, uma instituição estará apta a escolher a estratégia de preservação funcional mais adequada. Wilson (2007, p. 8) define as propriedades relevantes (*significant properties*) nos seguintes passos: "the characteristics of digital objects that must be preserved over time in order to ensure the continued accessibility, usability, and meaning of the objects, and their capacity to be accepted as evidence of what they purport to record." Essa definição foi testada para identificar propriedades relevantes em formatos de áudio, *email*, imagem *raster* e objetos de texto estruturado.

[21] Um exemplo de um registro central de formatos é o projeto *Digital Preservation Technical Registry*. Disponível na WWW: <http://ndha-wiki.natlib.govt.nz/current-initiatives/technical-registry>

[22] Um exemplo de política de recolhimento é a definida pelo Arquivo Nacional do Reino Unido. Disponível na WWW: http://www.nationalarchives.gov.uk/legal/takedown-policy.htm

Todavia, algumas considerações devem ser observadas quando se pretender o estabelecimento de propriedades relevantes, conforme argumentam Webb, Pearson e Koerbin (2013, p. 1):

> We have come to a tentative conclusion that recognising and taking action to maintain significant properties will be critical, but that the concept can be more of a stumbling block than a starting block, at least in the context of our own institution. We believe reference to significant properties in preservation planning requires some prior consideration of both the purposes for which digital content has been collected and the purposes of providing preservation attention. In effect, we are asking how can we know what attributes of digital materials we need to preserve if we haven't articulated why we are preserving them?

A instituição referida pelos autores é a Biblioteca Nacional da Austrália. Nesse artigo, eles apresentam as declarações de intenção de preservação acordadas naquela biblioteca, assim como os benefícios alcançados com esses acordos. Dessarte, esperam dialogar com outras instituições e com interessados no progresso da prática de planejamento da preservação e na exploração do modo com que tal planejamento orientado para as políticas pode ser efetivamente sistematizado.

2.5 − Metadados

Ao longo do processo da preservação digital, vários metadados deverão ser criados e mantidos. Os denominados metadados originais são criados por quem produz o objeto digital. Devem ser mantidos para assegurar a proveniência e autenticidade do objeto. Entretanto, outros tipos de metadado são acrescentados pelas instituições antes e depois da ingestão no repositório. Isto requer que a instituição defina uma política para os metadados, na qual se descrevem os tipos de metadado e os padrões[23] que são

[23] Jenn Riley elaborou um mapa visual com os 105 padrões de metadados mais difundidos na comunidade que cuida do patrimônio cultural até o ano de 2010. Disponível em WWW:http://www.dlib.indiana.edu/~jenlrile/metadatamap/

O *Digital Curation Center* reuniu os padrões por áreas de conhecimento: Biologia, Geociências, dados gerais de pesquisa, Ciências Físicas, Ciências Sociais e Humanas. Disponível em WWW:http://www.dcc.ac.uk/resources/metadata-standards

utilizados para certos tipos de objeto digital ou para uma coleção deles (Sierman *et al*, 2014).

O termo metadados,[24] no sentido que iremos utilizar nesta tese, reporta/significa um conjunto de informações estruturadas sobre um objeto digital. Os metadados agregam valor aos objetos que precisam ser preservados. Sua tipologia[25] é constituída por quatro categorias: descritiva, estrutural, técnica e administrativa.

Os metadados descritivos identificam a entidade intelectual por meio de propriedades, como autor e título, e ajudam na descoberta e acesso ao conteúdo dos objetos digitais. Esses metadados também podem resgatar o histórico de um objeto digital, isto é, eles podem determinar a proveniência do recurso digital. Dentre os padrões adotados para estruturar essa categoria de metadados, encontram-se o *Dublin Core Metada Initiative*, *MAchine-Readable Cataloging*, *Metada Object Description Schema* (Dappert e Enders, 2010).

Os metadados estruturais descrevem os relacionamentos de um objeto digital formado por vários arquivos físicos. Por exemplo: um livro pode ser estruturado em capítulos constituídos por um conjunto de páginas com cada uma delas em um arquivo digital distinto. Isto requer um registro dos relacionamentos entre os arquivos físicos e as páginas, entre as páginas e os capítulos e entre os capítulos e o livro todo. Geralmente, esses metadados são utilizados para a leitura realizada pelos computadores (Caplan, 2003; Gartner, 2008; Gartner e Lavoie, 2013). Um exemplo de padrão para esse tipo de descrição é o *Metadata Encoding & Transmission Standard* (METS).

Os metadados técnicos documentam as características dos objetos digitais detalhadamente. Por exemplo: é possível verificar se um arquivo no formato TIFF está fisicamente segmentado em ladrilhos ou tiras. Essa categoria de metadados complementa os metadados de preservação porque o conhecimento detalhado das características físicas de um objeto será determinante para reconstruí-lo ou migrá-lo para um outro formato.

[24] Caplan (2003) analisou a etimologia do termo metadado e concluiu que ele adquire diferentes significados de acordo com a comunidade e contexto em que ele for utilizado. A autora definiu metadado como uma informação estruturada sobre um recurso informacional em qualquer tipo de mídia ou formato. Neste caso, a comunidade é composta por bibliotecários e o contexto é o acervo físico e digital.

[25] A tipologia dos metadados que os identifica como administrativos, descritivos e estruturais foi descrita inicialmente em 1998 no MOA2 *White Paper* (precursor do padrão METS). Disponível em WWW: http://sunsite.berkeley.edu/moa2/wp-v2.pdf

Dois padrões exemplificam esse tipo de metadado: *Metadata for Images in XML Schema*; *Technical Metadata for Text*.

Os metadados administrativos contêm informações que facilitam o gerenciamento dos recursos. Podem incluir informações sobre o modo e o tempo em que um objeto foi criado, quem é o responsável pelo controle do acesso e armazenamento do conteúdo, quais são as restrições de acesso e uso, etc (Caplan, 2003). Segundo Dappert e Enders (2010), apesar dos demais tipos de metadado serem essenciais para a preservação digital, os administrativos são reiteradamente referenciados como metadados de preservação. Os autores analisaram os padrões de metadado na preservação digital e destacaram o *PREservation Metadata: Implementation Strategies* (PREMIS) como um exemplo de padrão elaborado especialmente para a preservação. Assim como o METS, o PREMIS é um padrão *de facto*, no que concordam Ruusallep *et al*. (2012), bem como Gartner e Lavoie (2013). Estes últimos, explicam o processo de empacotamento do PREMIS pelo METS e as complicações decorrentes dessa padronização.

Todas essas quatro categorias de metadado são desenvolvidas a partir do modelo de referência OAIS, que é um modelo de informação para os objetos gerenciados por um arquivo. O elemento central desse modelo é uma entidade denominada *information package*, que une, conceitualmente, o objeto e os metadados necessários à perenidade do seu uso. Segundo Day (2005), dentre os três pacotes de informação definidos no modelo OAIS, o *Archival Information Package* (AIP) é o mais importante a ser preservado. Não obstante, assim como o *Submission Information Package* e o *Dissemination Information Package*, um AIP se subdivide em dois tipos: *Content information*; *Preservation Description Information*. Ambos são encapsulados e identificados pelo *Packaging Information*, que, por sua vez, é localizável e recuperável pela *Descriptive Information* (figura 6).

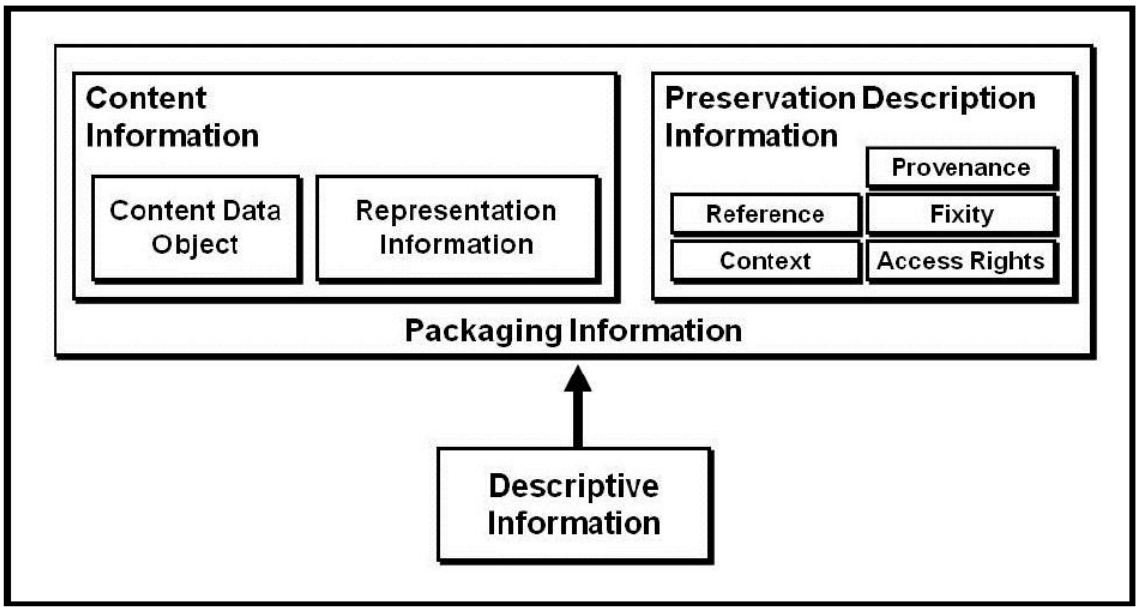

Figura 6 – The OAIS Archival Information Package

Fonte: Lavoie, 2014, p. 16

O elemento *Content Information* possui dois componentes: o *Content Data Object*, isto é, o objeto que se quer preservar (no caso dos materiais digitais, o objeto é uma sequência de *bits* - *bit stream*) e a representação da informação (*Representation Information*) requerida para tornar esse objeto inteligível aos usuários do OAIS. Assim, as categorias de metadado que satisfazem as características desse elemento são os metadados técnicos e os estruturais.

O elemento *Preservation Description Information* contém a informação que dará inteligibilidade ao *Content Information* ao longo do tempo. Esse elemento está focado na descrição dos estados passado e presente do *Content Information*, assegurando a este uma inequívoca identificação para que ele não seja alterado inadvertidamente (Day e Ross, 2005). O *Preservation Description Information* consiste nos seguintes componentes:

> - Reference Information uniquely identifies the Content Information within the OAIS's internal systems, as well as to entities and systems external to the OAIS. Examples include a system- generated internal identifier, and an ISBN.
> - Context Information describes the Content Information's relationships to other Content Information objects: for example, those that are related to it thematically (e.g., as part of a subject-based collection), or those that represent versions of the same content in alternative formats.
> - Provenance Information documents the history of the Content Information, including its creation, any alterations to its content or format over time, its chain of custody, any actions taken to preserve the Content

> Information (such as normalization or format migration), and the outcome of these actions.
> - Fixity Information ensures that the Content Information has not been altered in an undocumented way, through authenticity or integrity validation mechanisms such as checksums, digital signatures, or digital watermarks.
> - Access Rights Information documents any conditions or restrictions associated with the Content Information pertaining to both preservation and access. It may also include descriptions of rights enforcement mechanisms. Examples include licence terms, identification of those with authorized access permissions (e.g., a specified IP address range), and preservation terms and conditions negotiated between the OAIS archive and the Producer of the Content Information (Lavoie, 2014, p. 18).

O elemento *Packaging Information* faz a combinação de todos os componentes em um *Archival Information Package*, possibilitando ao usuário a identificação e localização de uma única unidade lógica dentro de um sistema de arquivo. Ele pode indicar o caminho de um diretório, nomes de arquivo ou pode ter a forma de um esquema de pacote mais detalhado como o METS.

O componente *Descriptive Information* é a informação que ajuda à descoberta e recuperação de um *Content Information*. Ele pode ser, por exemplo, um registro de metatado *Dublin Core*, derivado do *Content Information* associado ao seu respectivo *Preservation Descriptive Information*, mantido pelo OAIS para facilitar a descoberta a ser feita pelos usuários (Lavoie, 2014). As categorias que se inserem no *Preservation Descriptive Information* são os metadados descritivos e os administrativos.

2.6 – Propriedade intelectual

O material digital armazenado em RIs precisa ser regulado por contratos entre os *stakeholders* e protegidos por direitos de propriedade intelectual (*intellectual property*[26]) que não comprometam o seu livre acesso e preservação, pois ambas as ações implicam fazer cópias e redistribuí-las.

No ambiente das produções convencionais, o conceito de cópia é fundamental para o direito de propriedade intelectual, porque controla a reprodução de cópias onde

[26] In summary, "intellectual property" is a name that describes a class of several different legal regimes that generally concerns creations of the human mind. Copyright is but one of the legal regimes that fall under the umbrella of intellectual property. Disponível em WWW: http://corecopyright.org/2009/12/03/copyright_ip/

fizer sentido distinguir acesso e reprodução. Entretanto, no ambiente digital, o controle de cópia é um controle de acesso. Assim, os detentores de direitos e os usuários poderão ser afetados caso se aplique uma política que exagere o controle da reprodução do material digital (Borges, 2006). Doutra parte, a flexibilização dos direitos de propriedade intelectual nos RIs deve ser amparada pela observância da legislação do direito autoral vigente no país, especialmente no aspecto de cessão de direitos, como recomendam Dias, Sousa e Paiva (2012), no caso da preservação de periódicos científicos brasileiros de acesso aberto em uma rede LOCKSS. A lei de direitos autorais no Brasil nem sempre anda *pari passu* com as evoluções das TICs, o que requer, por exemplo, uma adequação de licenças do tipo *Creative Commons* às cessões de direitos autorais a fim de dar base legal, também, a um programa de preservação digital: "Na eventual utilização de uma Licença *Creative Commons* deve-se considerar que ela precisa estar em consonância com a cessão de direitos totais a título universal previsto na LDA [Lei de Direito Autoral] brasileira para permitir a replicação de conteúdos sem ferir os direitos de seus respectivos titulares." (Dias, Sousa e Paiva, 2012, p. 100).

Os contratos, licenças e passivos podem ser estabelecidos entre o repositório e os depositantes e entre o repositório e os prestadores de serviços. Devem ser elaborados de forma clara e mensurável, com definições de papéis, responsabilidades, prazos, condições, direitos e restrições de uso, além de ser disponibilizados facilmente para os interessados. Saliente-se que a formalização de contratos, licenças e passivos é uma das diretivas da norma ISO 16363 para a implementação de um repositório digital confiável (CTDE, 2014).

Satisfeitas as condições para o cumprimento dos requisitos legais, será necessário cuidar da preservação das informações de *copyright*, a fim de se evitar o problema dos *orphan works*, que são os trabalhos onde não é possível identificar os detentores de direitos autorais. Para tanto, Coley (2005) sugere a adição de elementos descritivos de *copyright* aos metadados dos trabalhos digitais ou digitalizados. Se os trabalhos forem removidos do seu ambiente original, correm o risco de perder alguns elementos que evidenciam o *status* de *copyright*. Por exemplo: a proveniência. A providência de elementos descritivos que podem ser transmitidos juntamente com o trabalho digital facilitaria o reuso do conteúdo intelectual que o trabalho representa. Assim, os metadados relacionados com o *copyright* deveriam ser considerados como elementos essenciais para

a descrição de recursos digitais e poderiam conter os seguintes elementos conceituais em sua estrutura:

> - General rights information: Copyright status (copyrighted, public domain, unknown); Publication status (published, unpublished); Dates (Year of copyright or creation, Year of renewal of copyright); Copyright statement (from the piece); Country of publication or creation.
> - Creator: Creator name, dates, and contact;
> - Copyright holder: Copyright holder contact;
> - Publisher: (Publisher name and contact); Year of publication;
> - Administrative data: Source of information (piece itself or other resources); Contact information; Rights research contact; Services contact (Coley, 2005, p. 7).

Tratando-se de instituições públicas que desejam utilizar um sistema *Digital Rights Management*[27] (DRM), os padrões de metadados de preservação recomendados são os seguintes: PREMIS,[28] Open Digital Rights Language,[29] METSRights,[30] XrML,[31] CopyrightMD,[32] MPEG21.[33] Entretanto, o sucesso da operação de um sistema DRM, assim como os cuidados com a preservação dos conteúdos e com os respectivos direitos digitais destes, dependem diretamente do investimento na qualificação da equipe responsável pelo RI (Kaur *et al.*, 2014).

Uma instituição também precisa estabelecer um acordo formal de depósito, de modo que deixe claro para o depositante e para a própria instituição os papéis e responsabilidades de ambas as partes. Um acordo pode ser geral ou específico, mas deve conter informações essenciais, tais como:

> - whether the organisation is allowed to delete material at all or only under certain circumstances;
> - the organisation's policy for disposal that identifies the material to be deleted and under which circumstances the material can be deleted by the organization;
> - the retention period for the material;
> - the measures needed to be in place to ensure that the deposit agreement is fulfilled;

---

[27] Um modelo de política *Digital Rights Management* foi desenvolvido para a plataforma de repositório *BlogForever*. Diponível na WWW: https://zenodo.org/record/7518

[28] Disponível na WWW: http://www.loc.gov/standards/premis/

[29] Disponível na WWW: https://www.w3.org/community/odrl/

[30] Disponível na WWW: http://www.loc.gov/standards/mets/

[31] Disponível na WWW: http://www.xrml.org/

[32] Disponível na WWW: http://www.cdlib.org/groups/rmg/

[33] Disponível na WWW: http://mpeg.chiariglione.org/

- whether the organisation is obliged to keep the original digital object, e.g. in case of migration or other kinds of preservation transformations;
- The information about keeping or disposing of the original in case of migration needs to be added to the information held about the digital object;
- access rights to the collection (Sierman *et al.*, 2014, p. 70).

Como vimos, no caso do Brasil, a legislação que regula a propriedade intelectual traz algumas dificuldades para as atividades de preservação digital e, mesmo nos países onde se criou acordos para adequar a preservação digital aos auspícios da lei o problema persiste. O *Digital Preservation Handbook* (DPH, 2015) defende uma negociação de direitos entre os repositórios, os depositantes e os detentores de direitos por meio de documentos formais, tais como, uma carta modelo para compensação de direitos do *staff*, um modelo de acordo para depósitos e um modelo de licenças. Considerando que diversas instituições estabeleceram esses tipos de modelos, o *handbook* lista os procedimentos mais usuais para fundamentar os modelos. Acrescenta, ainda, uma lista de verificação e um sumário das questões legais que devem ser consideradas nas licenças para preservação e acordos de depósito de materiais digitais, considerando as naturais diferenças entre as instituições e países.

Apesar dos vários tratados internacionais de propriedade intelectual, o desembaraço das nuances legais que dificultam a preservação digital só poderá ser levado a efeito se forem feitas modificações nas leis que regulamentam o *copyright* e os depósitos. Besek *et al.* (2008) revisaram a legislação de *copyright* e seu impacto nas atividades de preservação digital em quatro países: Austrália, Holanda, Reino Unido e Estados Unidos. Em conclusão, os autores recomendaram que os países deveriam estabelecer leis e políticas para estimular a preservação digital dos materiais protegidos por *copyright* que estão sob risco de se tornarem inacessíveis. Para tanto, eles propuseram uma lista de requisitos mínimos que essas leis e políticas deveriam possuir, sem prejudicar os direitos do autor.

2.7 − Padrões

Como as coleções digitais serão mantidas por diferenciadas equipes ao longo dos anos, é preciso assegurar a confiabilidade nas informações das coleções e nos padrões

que foram utilizados para criá-las e gerenciá-las. Por essa razão, o uso de padrões é altamente recomendado na preservação digital. Assim, uma instituição, quando declara seu firme propósito na adoção de padrões, está promovendo sua credibilidade (Sierman *et al*., 2014).

Os padrões representam acordos estabelecidos para formular as melhores práticas de preservação digital, embora tais práticas nem sempre sejam consensuais. Não obstante, a preservação digital depende da interoperabilidade entre os sistemas computacionais, o que torna a adoção de padrões uma ação imprescindível. Acrescente-se, ainda, que os padrões, além de funcionarem como ferramenta para se apoiar a acessibilidade, durabilidade e interoperabilidade das coleções digitais, possibilitam a troca e o uso de informações entre o *software* e o *hardware*. Todavia, a seleção, combinação e, eventualmente, a customização de padrões são desafios a serem enfrentados na adequação às necessidades de uma instituição (Ruusalepp *et al.*, 2012). Esses autores revisaram o estado da arte[34] do desenvolvimento de padrões e os problemas relacionados com a adoção de padrões, seja de ordem técnica ou institucional. No entanto, nesta seção, cumpre descrever os principais membros das famílias de padrões mais adotadas para a preservação digital, isto é:

- Os métodos de auditoria para repositórios digitais.
- A descrição de formato de arquivos.
- Os metadados de preservação constituem uma outra família de padrões, mas já foram abordados na seção 2.5.

A família de modelos de referência para repositórios digitais é formada pelos seguintes membros: *Open Archival Information Systems* (CCSDS, 2012), *Producer-Archive Interface Methodology Standard* - PAIMAS (ISO 20562, 2006), *Chain of Preservation Model* (InterPARES, 2007), *Digital Library Reference Model* (Candela *et al.*, 2011), *Curation Lifecycle Model* (DCC, 2015).

[34] O trabalho de Ruusallep *et al* (2012) é um dos capítulos do livro *Aligning National Aproaches to Digital Preservation* resultante de uma conferência internacional homônima, realizada na Estônia em 2011. A referida obra pretendeu contribuir para o progresso do campo da preservação digital ao longo dos 5 anos subsequentes à realização da publicação do livro, conforme se especula na conclusão da obra.

Na família de modelos de repositórios digitais, destacamos o Modelo de Referência OAIS porque este foi o primeiro padrão internacional a descrever um sistema de arquivamento digital. Assim, tornou-se o modelo básico para os demais padrões que surgiram na esteira da preservação digital e dos repositórios digitais, tais como o PAIMAS, TRAC e o dicionário PREMIS (Premis Editorial Committee, 2015).

Um OAIS é um arquivo (*archive*) constituído de uma organização (ou parte de uma organização maior), de pessoas e sistemas que aceitaram a incumbência de preservar a informação e torná-la disponível para uma comunidade de usuários (*Designated Community*). Entretanto, o termo *open* (aberto) se refere ao fato de as discussões sobre o desenvolvimento do modelo OAIS serem públicas. Não quer dizer que o acesso ao arquivo seja irrestrito (CCSDS, 2012). O uso do termo arquivo (*archive*), por sua vez, implica um sistema de arquivo dedicado à preservação da informação digital, tornando-a disponível a longo prazo. Implica, também, o cumprimento das seguintes responsabilidades, mesmo que este não seja assumido de forma absoluta:

> - Negotiate for and accept appropriate information from information producers;
> - Obtain sufficient control of the information in order to meet long-term preservation objectives;
> - Determine the scope of the archive's user community;
> - Ensure that the preserved information is independently understandable to the user community, in the sense that the information can be understood by users without the assistance of the information producer;
> - Follow documented policies and procedures to ensure the information is preserved against all reasonable contingencies, and that there are no ad hoc deletions.
> - Make the preserved information available to the user community, and enable dissemination of authenticated copies of the preserved information in its original form, or in a form traceable to the original (Lavoie, 2014, p. 8).

Os métodos de auditoria para repositórios digitais (*Digital Repository Audit Methods*) arrolados no trabalho de Ferreira *et al.* (2012) sobre o estado da arte em preservação digital inclui o *Trustworthy Repositories Audit & Certification: Criteria and*

Checklist (RLG-NARA, 2007), *Digital Repository Audit Method Based on Risk Assessment*, e o *Data Seal of Approval*.

A aplicação dos critérios do *Trustworthy Repositories Audit & Certification: Criteria and Checklist* (TRAC) toma por base o contexto da instituição, sua missão, prioridades e compromissos estabelecidos. Dessa forma, foram estabelecidos quatro princípios norteadores para a aplicar esses critérios:

> - **Documentation (evidence)**: The objectives, the design, specifications, and implementation of the digital long-term repository should be appropriately documented, and documentation should be reviewed and updated on a regular schedule. […].
> - **Transparency**: Ultimately, examining a repository for trustworthiness relies on another critical component: transparency, both internal and external. Only a repository that exposes its design, specifications, practices, policies, and procedures for risk analysis can be trusted. […].
> - **Adequacy**: The principle of adequacy takes into account that absolute standards do not exist for all aspects of repository organizational infrastructure, digital object management, and technologies and technical infrastructure. Even if they did, they would not apply to all types of repositories and archives and all situations. […];
> - **Measurability**: In principle, the goal is to have objective controls (criteria) against which repositories can be evaluated. […] (RLG-NARA - RLG, 2007, p. 7).

O TRAC originou a norma ISO 16363 (2012) – *Audit and Certification of Trustworthy Digital Respositories*. Segundo Carvalho *et al.* (2014), essa norma funciona como uma ferramenta para auditar, avaliar e certificar os repositórios digitais.

O *Digital Repository Audit Method Based on Risk Assessment* (DRAMBORA) é semelhante ao TRAC, mas ao invés de se centrar no repositório e sua plataforma tecnológica ele focaliza a identificação e a gestão de riscos. Para tanto, propõe os seguintes objetivos:

> - Elaborar um perfil organizacional, descrevendo e documentando a sua política de depósito, objetivos, responsabilidades, atividades e material custodiado;
> - Identificar e avaliar os riscos que poderão impedir a prossecução da sua missão e que ameaçam a salvaguarda dos seus materiais;
> - Gerir eficazmente os riscos, mitigando a sua probabilidade de ocorrência;
> - Estabelecer planos de contingência eficazes para minimizar os efeitos provocados por riscos que não puderam ser evitados (Ferreira, Saraiva e Rodrigues, 2012, p. 23).

As diretrizes de qualidade do *Data Seal of Approval* (DSA) destinam-se aos produtores de dados, instituições que geram dados digitais, organizações que arquivam dados e consumidores de dados. O DSA tem por objetivos proteger os dados, assegurar alta qualidade e orientar um confiável gerenciamento de dados para o futuro sem que seja necessário a implementação de novos padrões, normas ou altos custos. O selo (*seal*) atesta as seguintes garantias:

> - Gives data producers the assurance that their data and associated materials will be stored in a reliable manner and can be reused;
> - Provides funding bodies with the confidence that data will remain available for reuse and their investments will not be lost;
> - Enables data consumers to assess repositories where data are held;
> - Supports data repositories in the efficient archiving and distribution of data (DSA, 2014).

A descrição de formato de arquivo requer um entendimento profundo do conceito de formato de arquivo a fim de assegurar a preservação a longo prazo das coleções digitais. Um formato de arquivo é um conjunto de regras semânticas e sintáticas que controlam o mapeamento entre a informação e os *bits* que a representam (CDL, 2012). De acordo com Ruusalepp *et al.* (2012), o monitoramento da obsolescência dos formatos de arquivo é uma atividade chave na preservação digital, pois diversos projetos de preservação relataram uma certa dificuldade em se obter uma documentação e especificação confiáveis sobre os formatos. Isso fez surgir os repositórios de formatos sustentáveis para gerenciar a representação da informação de formatos, de modo que a informação esteja disponível para futuros curadores e preservacionistas.

O PRONOM[35] é um projeto do Arquivo Nacional do Reino Unido. Mantém um sistema *web* de informação sobre formatos de arquivo e seus respectivos *softwares* de suporte, além de desenvolver ferramentas e serviços de suporte às funções de preservação digital, tais como: avaliação de risco de preservação, planejamento de migração, identificação e validação de objeto e extração de metadados. Em seu portal, encontram-se disponíveis para livre acesso o sistema *web* PRONOM *Technical Registry*, a ferramenta de *software* livre *Digital Record Object Identification* (Esta ferramenta automatiza a identificação de formatos de arquivo em lote) e o serviço PRONOM *Unique Identifier*.

[35] Disponível na WWW: http://www.nationalarchives.gov.uk/aboutapps/pronom/tools.htm

O *Unified Digital Format Registry*[36] é uma base de conhecimento sobre formatos de arquivo, de livre acesso à comunidade de preservação digital. Trata-se de uma plataforma *open source* que incorpora as funções e propriedades do PRONOM e do *Global Digital Format Registry*.

O registrador *Digital Formats*[37] da Biblioteca do Congresso dos Estados Unidos disponibiliza informações sobre formatos de arquivo e suas implicações na preservação e sustentabilidade das coleções digitais.

O *File Format Risk Registry*[38] é um registrador de formatos criado pela *Open Preservation Foundation.* No entanto, ele se diferencia de outros registradores como o PRONOM e o *Unified Digital Format Registry*, porque seu foco está na resolução de problemas e na contenção de riscos. Dessa forma, esse registrador serve como um complemento aos registradores mais formais. Como é um *wiki*, ele não se amolda ao formalismo da abordagem centrada em base de dados.

Os modelos de dado desses registros evoluíram para uma adequação aos processos de planejamento de preservação. Assim, o trabalho progressivo nas propriedades significantes dos objetos (como o desenvolvido no projeto *Investigating Significant Properties of Eletronic Content Over Time*[39]), no desenvolvimento de linguagens (por exemplo: *Extensible Characterization Description Language*[40]) e ferramentas de planejamento de preservação (por exemplo: *Plato*[41]) fizeram avançar ainda mais a padronização da informação sobre formatos de arquivos digitais (Ruusalepp *et al.*, 2012).

2.8 − Acesso

A preservação digital, além de armazenar a informação, cuida de torná-la acessível e usável ao longo do tempo. Por conseguinte, as instituições precisam planejar

[36] Disponível na WWW: http://udfr.org/
[37] Disponível na WWW: http://www.digitalpreservation.gov/formats/index.shtml
[38] Disponível na WWW: http://wiki.opf-labs.org/display/TR/OPF+File+Format+Risk+Registry
[39] Disponível na WWW: http://www.significantproperties.org.uk/
[40] Disponível na WWW: http://www.planets-project.eu/docs/reports/Planets_PC2-D7_FinalXCDLSpec_Ext.pdf
[41] Disponível na WWW: http://www.ifs.tuwien.ac.at/dp/plato/intro/

o modo com que seus usuários irão acessar os objetos digitais. Isso inclui medidas simples, como a determinação do modo de um objeto ser visualizado, aplicação de métodos para se capacitar a comunidade de usuários para reusar e manejar os objetos digitais e habilitar sistemas para coletarem informações, como as que são fornecidas por meio dos metadados. Diversas abordagens poderão ser escolhidas de acordo com o material do acervo. Por exemplo: o acesso a um *website* pode usar uma abordagem distinta do acesso a um conjunto de dados.

A usabilidade da coleção digital preservada deve ser assegurada pela instituição mantenedora de um acervo. Para tanto, ela deve dominar o conhecimento das características significativas dos objetos digitais. São elas:

- Os formatos de arquivo;
- As características do objeto;
- O ambiente (*software/hardware*) necessário para se apresentar o objeto;
- A comunidade alvo e suas respectivas necessidades (Sierman *et al*, 2014).

Quando um objeto digital sofre ações de preservação no seu ambiente nativo, um novo objeto é criado e, por sua vez, é apresentado ou executado em um novo ambiente. Por exemplo: um arquivo do Word pode migrar do seu editor *Microsoft* para um editor *Adobe* em formato PDF. Naturalmente, haverá o risco de perda ou alteração das características do objeto original. Neste caso, poderá ocorrer a perda das macros originais, do histórico de edições e de um grau de interatividade não suportado pelo PDF.

As características significativas são uma forma de metadados de preservação. Refletem os requisitos de um negócio e capturam as características de um objeto original e o ambiente dele, por meio de uma ação de preservação. Por exemplo: alguém determina que, em uma coleção de jornais, todas as páginas mantenham suas margens originais. Essa determinação irá influenciar a decisão sobre quais ações de preservação deverão ser adotadas. Diversos projetos se empenharam em procurar soluções para essas ações: o PREMIS dá suporte à captura de propriedades significativas de objetos digitais; o registro de formatos PRONOM trabalha na identificação de propriedades aplicáveis aos formatos de arquivos; o projeto InSPECT trabalha na identificação de propriedades aplicáveis ao tipo de conteúdo, tais como imagens e *e-mails*; e a *Open Preservation Foundation*

pesquisa as características significativas avançadas e usam-nas no planejamento de preservação (Dappert e Enders, 2010).

O acesso à informação digital está condicionado pelos direitos digitais. Estes são previstos pelo modelo de informação OAIS (*Access Rights* – figura 6). Conforme o verbete *digital rights* da *Wikipedia*, o termo direitos digitais descreve os direitos individuais para acessar, utilizar, criar e publicar conteúdos digitais, acessar e usar computadores ou outros dispositivos eletrônicos e redes de comunicação. Na interpretação de Kaur *et al.* (2014), isto significa que os indivíduos podem reutilizar os conteúdos digitais em dispositivos eletrônicos.

Os direitos digitais podem ser incorporados no próprio objeto digital (por exemplo, por meio do uso de senhas ou metadados) como resultado de um acordo com o produtor de uma coleção ou baseado em alguma lei nacional ou internacional. Ao longo dos anos os direitos digitais poderão sofrer alterações decorrentes de um contexto de mudanças. Assim, a fim de preservar sua reputação, uma instituição precisa definir uma política que enumere os direitos importantes e descreva como a instituição irá tratá-los. Não obstante, os direitos de acesso podem ser aplicados apenas a uma parte da comunidade de usuários (por exemplo: usuários de uma determinada coleção) ou a uma coleção específica (Sierman *et al*, 2014).

No modelo funcional do OAIS, a entidade denominada acesso (*access*) gerencia os processos e serviços, pelos quais os usuários e a comunidade alvo (*Consumers; Designated Community*) localizam, requisitam e recuperam os objetos armazenados. Essa entidade funcional, uma vez que representa a interface do OAIS com seus usuários, é o mecanismo primário, pelo qual, o OAIS assume a responsabilidade de tornar disponível à comunidade de usuários o material armazenado. Quando uma requisição de acesso é solicitada, um pacote de informações denominado *Dissemination Information Package* (DIP) é entregue ao usuário. O DIP é um dos três componentes do modelo de informação OAIS. O conceito do DIP enfatiza o fato de que o pacote de informações entregue ao usuário pelo OAIS pode diferir na forma e conteúdo daquele que está armazenado no acervo. Entre as inúmeras diferenças entre um DIP e um AIP estão aquelas relacionadas com o formato de arquivo (uma imagem TIFF pode ser convertida em JPEG antes da entrega); a quantidade de conteúdo (um DIP pode corresponder a um ou vários AIP e até

mesmo só parte de um AIP) e a quantidade de metadados fornecidos com o conteúdo (existe a possibilidade de um DIP não conter todos os metadados associados a um objeto digital, visto que a maioria deles pode não interessar ao usuário ou *consumer*) (Lavoie, 2014).

Uma vez que os usuários e as comunidades-alvo diferem nas necessidades de uso dos objetos, então é recomendável que uma instituição desenvolva uma política cuja abordagem reflita os interesses dos usuários. Para estes, a usabilidade pode estar atrelada ao modo de ser apresentado o material: alguns podem ficar satisfeitos com a apresentação contemporânea do objeto digital, enquanto outros podem necessitar de uma apresentação com múltiplos objetos. Do mesmo modo, a usabilidade está relacionada com a garantia do acesso ao longo do tempo e com a condição que estabelece a possibilidade para uma versão original ou uma derivada ser apresentada. Uma instituição, se deseja suprir as necessidades da comunidade de usuários, deverá considerar o desenvolvimento de DIPs para sustentar uma política de preservação.

Na seção anterior, vimos que uma das responsabilidades atribuídas a um arquivo do tipo OAIS é a de assegurar que a informação preservada será compreendida, de maneira autônoma (*independently understandable*), pela comunidade de usuários, isto é, os usuários poderão entender a informação sem precisar recorrer aos produtores. Assim, Sierman *et al* (2014) recomendam que sejam acrescentadas explicações para manter a informação no objeto digital inteligível no futuro. Para tanto, se faz necessário implementar uma política que descreva as ações pretendidas por uma instituição a fim de possibilitar um entendimento autônomo por parte dos usuários. Um exemplo desse tipo de responsabilidade é o que foi estabelecido no item *Operating Principles*, *Section A*, da política de preservação digital da Universidade de Utha, onde se lê:

> *The Library will strive to:*
>
> - *Comply with OAIS and other digital preservation standards and practices;*
> - *Ensure that content remains readable and understandable;*
> - *[…]* (JWLB, 2012, p. 4).

Na terminologia do Modelo de Referência OAIS, o termo *access aid* designa um *software* ou documento que permite os usuários (*consumers*) localizar, analisar, ordenar ou recuperar uma informação de um arquivo do tipo OAIS. O termo *finding aid* é um tipo

de *access aid* que permite um usuário pesquisar e identificar um *Archival Information Package* (AIP) que lhe interessar (CCSDS, 2012). Isto significa que uma instituição deve levar em conta o modo com que os usuários desejam descobrir uma informação no repositório, o que pode ser feito com o uso de metadados descritivos ou de identificadores persistentes que já vierem com o objeto digital. A ausência de uma política que esclareça como a informação digital será acessada poderia comprometer a demanda do uso das coleções em um repositório digital (Sierman *et al*, 2014).

O OAIS define comunidade-alvo (*Designated Community*) como um grupo de usuários potenciais que deveriam estar aptos a entender um conjunto de informações específicas. A comunidade-alvo pode ser constituída de múltiplas comunidades de usuários. Ela é definida pelo repositório e essa definição pode mudar ao longo do tempo (CCSDS, 2012). Conhecer a comunidade-alvo influi na qualidade do serviço de um repositório e torna usável e inteligível, a longo prazo, a informação preservada. Essa comunidade pode ser constituída por usuários externos e internos, inclusive por outras instituições. Ela pode diferir por coleções e assim ser definida pelo tipo de coleção (Sierman *et al*, 2014). Um exemplo de política que define claramente sua comunidade-alvo é encontrada na seção *Access and Use* da política de preservação digital do *Inter-university Consortium for Political and Social Research – Michigan University*, onde se lê:

> The designated community at ICPSR, as described by OAIS, includes traditional users, i.e., social science researchers and graduate students at member institutions; and newer categories of users, e.g., undergraduates, policymakers, practitioners, and journalists. To protect the identity of human subjects who may be represented in the deposited data, ICPSR devotes significant resources to developing and implementing the means to ensure confidentiality (ICPSR, 2012, p. 1).

2.9 − Organização

As soluções para a preservação digital vão além das questões de ordem técnica: elas também são de ordem organizacional. A preservação digital implica a interação entre o ambiente de preservação e as questões de procedimentos e objetivos organizacionais mais amplos. Por exemplo: questões relacionadas com pessoal, gestão de riscos,

orçamentos, custos,[42] funções e responsabilidades. Nesse sentido, torna-se imperativo que os membros de uma instituição reconheçam o valor dos dados digitais como um elemento crucial para o sucesso do funcionamento da instituição (Beagrie *et al.*, 2008).

Um estudo realizado pelo *Digital Curator Vocational Education Europe Project*, em 44 países, entre o ano de 2011 e o de 2012, constatou uma acentuada carência de pessoal qualificado para as atividades de preservação digital e de opções de formação adequada nesse campo do conhecimento. Segundo Engelhardt (2013), autora desse estudo, o pessoal que trabalha com preservação digital precisa ter um largo espectro de habilidades e competências. Tal espectro compreende as competências e habilidades técnicas genéricas e o domínio do conhecimento técnico específico da preservação digital; no entanto, os participantes da pesquisa manifestaram como a necessidade mais urgente o conhecimento especializado da preservação digital, sobressaindo-se a demanda pela formação básica em preservação digital e em planejamento de gestão de dados.[43]

O projeto DigCurV é o autor do *DigCurv Curriculum Framework*, que é um meio para identificar, avaliar e planejar os treinamentos requeridos para habilitar o pessoal engajado na curadoria digital, tanto no presente como no futuro (DigCurv, 2013). A formação de uma equipe especializada é fundamental para o sucesso de um programa de preservação digital a longo prazo. Uma política de preservação digital deverá declarar o interesse da instituição em investir em recursos humanos.

No ano de 2012, uma pesquisa conduzida pelo *Standards and Practices Working Group* da *National Digital Stewardship Alliance* (NDSA) investigou 85 instituições, sendo a maioria norte-americanas, com mandatos de preservação de conteúdo digital, para se saber como elas formaram e organizaram suas funções de preservação. Dentre as várias descobertas, destacamos aquelas que revelaram o desejo das organizações de duplicar a quantidade de membros de suas equipes, dedicados às atividades de preservação digital, os repetidos treinamentos dos grupos existentes e os pré-requisitos desejáveis para os novos gestores dedicados à preservação digital, quais sejam: ser

[42] O portal *Curation Costs Exchange*, lançado em 2014, fornece orientações sobre custos com a curadoria digital. Disponível na WWW: http://www.curationexchange.org/understand-your-costs

[43] O *Northeast Document Conservation Center* propõe um *curriculum* para um curso introdutório de preservação e disponibiliza os planejamentos de aula no seu *website*. O planejamento de número 11 é feito para aula de preservação digital. Disponível na WWW: https://www.nedcc.org/curriculum/lesson.introduction.php .

apaixonado e motivado pela preservação digital, assim como ter um conhecimento sobre os padrões, as melhores práticas e ferramentas de preservação digital. Os autores dessa pesquisa, Atkins *et al* (2013), recomendam estudos comparativos com os trabalhos desenvolvidos por Claudia Engeharldt[44], Ruben Riestra[45], Naomi Nelson[46] e Jeonghyun Kim.[47]

Uma equipe técnica é apenas uma das partes interessadas (*stakeholders*) na preservação digital, conforme está resumido no quadro 4.

Quadro 4 – O interesse e envolvimento dos *stakeholders* na preservação digital

Type of institution	Interest
Cultural heritage institutions	Management: informed decisions on preservation Curators: understanding and efficient use of tools/services
Institutions that design policies on national and international level	Monitoring, guidance, evaluation of progress
Academic and research institutions	Use digital objects for research and teaching that requires long-term preservation
e-Infrastructures	Preserve distributed resources and provision of services to a variety of stakeholders
R&D institutions active in digital preservation	Develop new solutions in digital preservation
Auditors	Perform evaluation of a number of aspects of digital preservation infrastructures, workflows, policies and procedures
Fund-makers	European Commission and other funding bodies that fund research and development in digital preservation and through this further the state of the art
End-users	End users or consumers are the beneficiaries of reliable digital preservation but this does not mean that they necessarily understand the domain and its issues in detail

Fonte: Ruusalepp e Dobreva (2012, p. 30)

A análise de Ruusalepp e Dobreva (2012) leva em consideração dois pontos de vista sobre os *stakeholders*: a visão centrada no usuário e a centrada nos dados. Nesta seção, importa explicar o primeiro ponto de vista. A comunidade de usuários é formada por um certo número de usuários com diferentes demandas.

---

[44] *Report and analysis of the survey of training needs*. 2012. Disponível na WWW: http://www.digcur-education.org/eng/Resources/Report-and-analysis-on-the-training-needs-survey

[45] *D36.1 – Business Preparedness Report*. 2013. Disponível na WWW:http://www.alliancepermanentaccess.org/wp-content/uploads/downloads/2013/03/APARSEN-REP-D36_1-01-1_0.pdf

[46] *SPEC Survey on Managing Born Digital Special Collections and Archival Materials.* 2012. Disponível em WWW: http://publications.arl.org/Managing-Born-Digital-Special-Collections-and-Archival-Materials-SPEC-Kit-329/

[47] *Digital curation in the academic library job market*. 2012. Disponível na WWW:<https://www.asis.org/asist2012/proceedings/Submissions/283.pdf>

Primeiro, os provedores de conteúdo precisam preservá-los. Eles são de vários tipos: instituições nacionais com alta produção de conteúdos digitais (por exemplo: bibliotecas nacionais e seus conteúdos digitalizados); pequenas instituições com orçamentos limitados, mas estas necessitam resolver seus problemas técnicos e de planejamento na área da preservação digital; editoras e outros criadores de conteúdo que têm a obrigação legal de preservar o material digital que produzem.

Segundo, os planejadores de políticas e os detentores de programas em diferenciados âmbitos territoriais (nacional, regional e local) investem no desenvolvimento de infraestrutura de pesquisa e precisam de orientação ao longo do tempo.

Terceiro, a comunidade de usuários finais precisa da infraestrutura para assegurar o acesso aos objetos digitais de suas pesquisas ou de seus consumos culturais.

Quarto, a comunidade de desenvolvimento e pesquisa, que atua no domínio da preservação digital, adotaria soluções ou as desenvolveria.

Ruusalepp e Dobreva (2012) afirmam que o número de especialistas na área de preservação digital cresceu nos últimos anos; existe, porém, um claro panorama a indicar a necessidade de especialistas com formação acadêmica e treinamento profissional.

Um outro aspecto organizacional a ser considerado em uma política de preservação digital é a gestão de riscos. Segundo o glossário do DRAMBORA (McHugh *et al.*, 2007), a gestão de riscos é o conjunto de atividades coordenadas para dirigir e controlar uma organização, no que diz respeito aos riscos. No caso da preservação digital, isso implica identificar e mitigar os riscos inerentes ao ambiente tecnológico e às coleções. Essas ações devem fazer parte da rotina de atividades relacionadas com a preservação amparadas por procedimentos adequados e pessoal qualificado. Entretanto, Sierman *et al.* (2014) ressalvam que as medidas de contenção de riscos podem ser aplicadas de modo diferenciado sobre uma mesma coleção ou coleções, mas o critério seletivo vai depender da importância das coleções e dos recursos disponíveis. Por outro lado, advertem que uma instituição, se não se conscientizar sobre os riscos, poderá perder suas coleções e, consequentemente, frustrar seus objetivos e comprometer a sua própria reputação.

Um programa de preservação digital tem um custo que deve ser previsto no orçamento da instituição, a fim de dar sustentação ao programa. O interesse nestes temas deverá ser declarado em uma política de preservação digital. O cerne desse programa é o RI. Conforme ensinam Burns, Lana e Budd (2013), a administração de um RI requer um compromisso a longo prazo para salvaguardar, preservar e tornar acessível o conteúdo intelectual do acervo digital. Se uma instituição não compreende o significado desse serviço, o valor de um tal programa pode ser subestimado. Por conseguinte, os investimentos para manter a sobrevivência e o crescimento de um RI minguariam. Por essa razão, Lynch (2003) adverte que é fácil criar um serviço de RI, mas é difícil mantê-lo. Portanto, uma instituição deveria pensar bem antes de lançar um RI.

A preocupação com as funções e responsabilidades do pessoal envolvido com um programa de preservação digital complementam o elenco dos aspectos organizacionais discutidos nesta seção. Para Sierman *et al.* (2014), uma instituição deveria deixar claro para seus servidores quem assume determinadas funções e a quem competem certas responsabilidades, inclusive registrando estes aspectos nos procedimentos que são atualizados regularmente. Durante o ciclo de vida das coleções digitais, várias tomadas de decisões e ações de planejamento irão determinar as coleções que serão criadas, os critérios sob os quais elas serão aceitas, as medidas de controle de qualidade e a aprovação dos planos de preservação. Dessa forma, o sucesso de um programa de preservação digital depende, também, do conhecimento que a instituição tem sobre o pessoal envolvido e quem está autorizado a tomar decisões. Do contrário, a falta de uma definição das funções e responsabilidades do pessoal envolvido com a salvaguarda das coleções pode contribuir para decisões que não estão amparadas pela política institucional de preservação digital.

2.10 − Auditoria e certificação

Um repositório digital confiável é aquele que demonstra ter consciência dos riscos e ameaças inerentes a sua natureza. Neste sentido, ele demonstrará seu cuidado com os materiais depositados por meio de um constante monitoramento, planejamento e implementação de estratégias que deem sustentação a sua missão de preservação. Todas essas ações implicam um empreendimento complexo e caro, que os depositantes,

stakeholders, financiadores e a comunidade-alvo, além de outros repositórios digitais, terão de se unir, necessariamente, a uma rede de colaboração mais ampla de modo que preservem uma vasta quantidade de informações digitais, a qual é criada no presente e que será no futuro. A comunicação dos resultados de auditagem sinaliza a transparência necessária para dar credibilidade ao repositório e promove uma confiança ainda maior tanto no próprio repositório como no sistema que o integra. Mesmo assim, o *status* de confiabilidade só será alcançado se as auditorias e a obtenção de certificações (nem sempre imprescindíveis) forem realizadas periodicamente (ISO 16363, 2012). Segundo Sierman *et al*. (2014), os principais padrões de auditoria e certificação são: *Data Seal of Aproval*, ISO 16363, DIN 31644. Os dois primeiros foram explicados de forma introdutória na seção sobre padrões, neste capítulo.

O DIN 31644 – *Criteria for Trustworthy Digital Archives* – foi publicado em 2012. Está modelado de acordo com o OAIS e o seu glossário adota conceitos do dicionário PREMIS (Premis Editorial Committee, 2015). Um padrão como o DIN 31644 serve para indicar sinais de confiabilidade a seus usuários e atestar a credibilidade pública do repositório. Simultaneamente, a auditagem, com base nesse padrão, pode revelar possíveis pontos de fragilidade no repositório. Por outro lado, serve para encorajar o investimento dos financiadores públicos ou privados. O ciclo de vida de um repositório digital inicia-se pelo planejamento seguido pela especificação e implementação, acrescido de, no mínimo, intenções de melhorias, o que recomenda a elaboração de um novo planejamento. Os 34 critérios desse padrão estão fundamentados em quatro princípios, que são os mesmos do TRAC, como foi visto na seção sobre padrões, neste capítulo: documentação (registro das atividades), transparência (ao menos uma parte da documentação deveria estar disponível para os usuários, financiadores e público), adequação (as soluções para o arquivo devem ser adequadas aos seus objetivos e tarefas) e mensuração (o critério deveria ser tanto quanto possível concreto e mensurável) (Keitel, 2012).

No âmbito da União Europeia, em julho de 2010, os grupos de trabalho do *Repository Audit and Certification, Data Seal of Approval* e do *Deutsches Institut für Normung* assinaram um memorando de acordo acerca de um *framework* integrado para auditar e certificar repositórios digitais. O *framework* consiste em uma sequência de três níveis que incrementam a confiabilidade de um repositório digital:

- Certificação Básica – Este nível é dado aos repositórios que obtiverem a certificação DSA;
- Certificação Estendida[48] – Este é dado aos repositórios que possuem a Certificação Básica e realizaram uma auditoria interna de forma estruturada, revisada externamente e disponibilizada ao público, com base na ISO 16363 ou DIN 31644;
- Certificação Formal – Este certificado é concedido aos repositórios que, em adição à Certificação Básica, obtiveram uma auditoria externa completa e uma certificação baseada na ISO 16363 ou DIN 31644 (Giarreta *et al.*, 2010).

Em Portugal, por exemplo, no âmbito do projeto RCAAP, no ano de 2014, foi realizada uma auditoria nos repositórios do serviço SARI, com base na ISO 16363,[49] pretendendo-se o aumento da confiança dos usuários e a transparência das soluções adotadas, de modo que se elevasse o nível de maturação desses repositórios em três dimensões:

> **1. Infraestrutura organizacional**
>
> Esta dimensão abrange a estrutura governativa do repositório e a sua viabilidade organizacional, analisando as questões relacionadas com a gestão do repositório, processos e recursos humanos afetos. Inclui ainda: políticas de preservação, documentação dos processos, sustentabilidade financeira das instituições que gerem os repositórios, contratos, licenças e responsabilidades do serviço.
>
> **2. Gestão de objetos digitais**
>
> Esta dimensão analisa o processo de ingestão e gestão de objetos digitais do repositório, ou seja, a forma como se incorpora a informação digital, assim como a criação e gestão de um *Archival Information Package* (AIP). Engloba ainda o planeamento da preservação e a forma como um AIP será preservado. Finalmente, observa as componentes de gestão de informação do serviço e a gestão de acessos.
>
> **3. Infraestrutura e gestão da segurança**
>
> A última dimensão engloba as questões técnicas relacionadas com a gestão e controlo de riscos inerentes à infraestrutura e à gestão da segurança. Esta componente relaciona-se diretamente com as infraestruturas disponibilizadas pelo serviço SARI do projeto RCAAP (Carvalho *et al.*, 2014, p. 5).

[48] O selo NESTOR pode garantir uma certificação estendida. Disponível na WWW:http://www.langzeitarchivierung.de/Subsites/nestor/EN/nestor-Siegel/siegel_node.htmltml

[49] Para saber como prepara uma auditoria com a norma ISO 16363, consulte o *website* do *Primary Trustworthy Digital Repository Authorisation Body*. Diponível na WWW: http://www.iso16363.org/preparing-for-an-audit/

O relatório de auditoria concluiu que, em média, os repositórios do SARI estavam em um nível incipiente de maturidade; uma nova auditoria, porém, foi prevista a fim de averiguar se os repositórios envidaram os esforços necessários para atender as recomendações sugeridas pelo próprio relatório. Em outros termos: seria realizada uma nova auditoria para reavaliar o nível de maturidade dos repositórios. Apesar de o relatório não mencionar qualquer intenção de certificar os repositórios nos moldes do memorando publicado no *website Trusted Digital Repository*, a divulgação da auditoria e a transparência dos seus métodos já são indicações evidentes de quanto podemos confiar no SARI. Resta confirmar se essa divulgação está sendo feita também nos próprios *websites* dos repositórios assistidos pelo serviço.

Não se podem cobrir todas as situações possíveis nas métricas nem precisar tudo quanto um repositório deve fazer por obrigação. Isto se aplica a todos os tipos de auditoria. Assim, boa parte dos problemas deve ficar ao encargo do julgamento dos auditores. Para tanto, a melhor forma de entender como a norma ISO 16363 foi escrita é ter em mente que os autores da norma se preocuparam de quão bem um repositório preserva a informação digital que lhe foi confiada. Isto significa que a pista inicial para orientar os auditores está na seguinte declaração: "Well at least look at the organisation – make sure it cannot suddenly go out of business, and also make sure that they know how to preserve the digital objects" (Giarreta, 2011, p. 465). Sendo assim, pode-se dizer que há duas recomendações para os auditores:

- Considerar a organização e suas finanças;
- Considerar a maneira que a organização cuida do material digital.

Na verdade, existe um complemento para a segunda recomendação, que tem sido ignorado de um modo geral: certificar-se de que os bens digitais não podem ser furtados ou perdidos. Nesta assertiva, está implícita a sugestão de uma auditoria exclusiva para os aspectos de segurança. Esta auditoria levaria em conta um grupo de métricas adequados para uma certificação ISO 27000. Giarreta (2011) critica o fato de uma tal certificação não estar sendo exigida de forma alguma. Todavia, Houghton (2015) assegura que a norma ISO 16363 (2012) é, de certo modo, a mais largamente aceita, especialmente no contexto dos repositórios institucionais.

Capítulo 3 − Modelos de política de preservação digital

Uma política é um conjunto de regras ou princípios que orientam as tomadas de decisão e as ações para se alcançarem os resultados desejados relativos a um determinado aspecto ou objetivo. Não deve ser prescritiva. Deve receber a chancela do nível mais alto da hierarquia organizacional, ser tecnologicamente neutra e dar suporte à estrutura de governança e à cultura organizacional. Serve como um modelo que delineia o escopo e os requisitos dos procedimentos para a prática da preservação digital (InterPARES, 2012). O conjunto dos procedimentos práticos constitui um programa de preservação digital.

A preservação digital é um problema de pesquisa estudado há décadas. Várias tecnologias foram desenvolvidas para evitar a perda do material digital produzido pelas instituições. Apesar disso, uma grande parte delas não possuem uma cultura de preservação consolidada. Conforme Lyman e Bresser (2010), a preservação digital a longo prazo requer não apenas soluções técnicas e estratégias organizacionais, mas também a formação de uma nova mentalidade de valorização e apoio à sobrevivência dos *bits* no transcurso do tempo. Muitos estudos seguem nessa direção. Destarte, o passo inicial está na elaboração de uma política de preservação digital.

A base de um programa de preservação digital é construída por uma política de preservação digital, porquanto esta política fornece um fundamento intelectual sólido e consistente para as soluções práticas. Tal política também pode assegurar o envolvimento de toda uma organização ou instituição com os princípios e práticas desse tipo de preservação. A implementação de um tal programa deve ser coerente com a política estabelecida. Esta, por sua vez, deve ser declarada por escrito a fim de externar o compromisso da organização ou instituição com a preservação do seu material digital (Brown, 2013; ERPANET, 2003).

No trabalho desenvolvido por Beagrie *et al.* (2008), para ajudar as instituições de ensino superior a desenvolverem e implementarem suas políticas de preservação digital, concluiu-se que toda política desse tipo deverá refletir as diretrizes e estratégias de uma instituição. Ao longo de anos, as universidades tinham como diretriz o aproveitamento do conteúdo digital e o dos serviços eletrônicos, o que resultava em consideráveis

benefícios de flexibilidade e ganho de produtividade. Mas, em meados da década de 2010, a prioridade foi mudando para o desenvolvimento de estratégias e infraestrutura de suporte, a fim de garantir o acesso ao conteúdo digital e, assim, manter os benefícios alcançados em anos anteriores. Entretanto, os autores advertem que todo acesso a longo prazo e todo benefício futuro estarão fortemente condicionados à adoção de estratégias de preservação apoiadas em uma política de preservação proeminente.

A elaboração de uma política deve assentar em seções predefinidas, de modo que fique clara a sua abrangência e os seus limites. Autores, como Brown (2013) e Noonan (2014), analisaram as seções mais encontradas nas políticas de preservação digital de organizações e instituições. Neste sentido, diversas iniciativas nacionais e projetos multinacionais orientam a elaboração, o planejamento e a implementação de uma política de preservação digital disponibilizando seus respectivos modelos de *framework* livremente na *Web*. Entre eles, contam-se o *Joint Information Systems Committee* (JISC), o *Scalable Preservation Environments* (SCAPE), *International Research on Permanent Authentic Records in Electronic Systems* (InterPARES).

3.1 − Ferramenta de política do Directory of Open Access Repositories

O *Directory of Open Access Repositories* (OpenDOAR) é um serviço financiado pelo JISC. Assegura uma lista qualificada de repositórios de acesso aberto do mundo todo. As indicações do OpenDOAR são importantes, porque se trata de um ponto de visibilidade mundial para os repositórios de acesso aberto, com critérios de tecnicidade e acesso reconhecidos pela comunidade do movimento do acesso aberto. Um repositório, quando é cadastrado no OpenDOAR, tem suas informações avaliadas, na consistência e qualidade, por um *staff* de especialistas. O serviço disponibiliza uma ferramenta de políticas, a qual procura suprir um mínimo de recomendações em conformidade com o movimento de acesso aberto. As políticas abordadas são:

a) Política de metadados: informações que descrevem os itens nos repositórios; acesso e reuso dos metadados.

b) Política de dados: para textos e outros dados completos. Acesso e reuso dos dados na íntegra.
c) Política de conteúdo: para tipos de documento e conjunto de dados; tipos de repositório; tipos de material armazenado; principais idiomas.
d) Política de submissão: relacionada com os depositantes, com a qualidade e com o *copyright*; depositantes elegíveis; regras de depósito; moderação; controle de qualidade do conteúdo; embargos dos editores e financiadores; política de *copyright*.
e) Política de preservação: período de guarda; preservação funcional; preservação do arquivo; política de remoção; itens removidos; controle de versão; política de fechamento (OPENDOAR, 2014).

Na verdade, todas essas políticas compõem uma política de preservação digital; o OpenDOAR, porém, especificou uma política exclusivamente para a preservação e a denominou de política de preservação. Suas seções são resumidas a seguir:

- O período de guarda estabelece opções para se determinar por quanto tempo o objeto será retido ou verificar se esse tempo é indefinido.
- A preservação funcional estabelece a hipótese de a política tentar assegurar a legibilidade e acessibilidade dos itens armazenados, a forma com que fará isso e o modo com que o repositório está trabalhando com parceiros externos, se for este o caso.
- A preservação de arquivo (o item armazenado no repositório) poderá ser feita por meio de *backups* regulares, preservação do *bitstream* e microfilmagem. Assim como na seção anterior, as opções não são exclusivas.
- A política de remoção pode ser definida pela escolha de uma entre três opções (indefinida, a critério do repositório ou requerida por detentores de direitos autorais). Os motivos para uma remoção podem ser definidos por seis opções não exclusivas.
- A seção itens removidos oferece três opções para se estabelecer como os itens serão removidos e diversas opções para se definir como os identificadores serão tratados.

- O controle de versão é realizado em seis opções não exclusivas e trata do que pode ser feito com as versões armazenadas.
- A política de fechamento trata do possível fechamento de um repositório, do que será feito com a base de dados e decide a detenção ou a devolução dos itens aos depositantes.

A despeito de pretender o reconhecimento como um ponto de referência global para os repositórios de acesso aberto, o OpenDOAR não avançou no desenvolvimento de sua ferramenta. Ela foi criada em 2006 e até novembro de 2016 não havia sido atualizada. Na verdade, trata-se de um roteiro limitado e defasado.

A seção sobre o período de retenção está fora de contexto, porque não se sabe quais são os critérios que determinam o tempo de armazenamento dos objetos e porque os objetos seriam removidos, pois mantê-los acessíveis a longo prazo é uma premissa. A seção de preservação funcional não leva em conta algumas condições a serem observadas antes de se poder adotá-la: mudanças tecnológicas, riscos, demandas de uma comunidade específica e uso de padrões. A seção de preservação de arquivo (objeto digital) também não leva em conta a complexidade envolvida na preservação de *bit*. Há duas seções dedicadas à remoção de objetos armazenados no repositório e uma incompreensível previsão de desativação do repositório. Uma instituição, se aderisse a esse modelo, não precisaria fazer controle de versão, como é denominada a última seção. Curiosamente, observa-se que nenhum dos 26 repositórios institucionais das universidades federais do Brasil, cadastrados no diretório, preencheu as políticas apresentadas pela ferramenta.

Essa ferramenta de política não poderia servir como um instrumento mandatório, pois não tem um arcabouço político verificável. Ela não prevê o envolvimento da administração superior de uma instituição nem o dos *stakeholders*. Não estabelece cláusulas de princípios e implementação nem tampouco estabelece um direcionamento, visando à implementação de procedimentos práticos para manter um objeto digital autêntico e acessível ao longo do tempo.

3.2 − O modelo de *framework* de política do JISC

Beagrie *et al*. (2008) elaboraram um modelo delineador de políticas institucionais de preservação digital para ser aplicado em instituições de ensino superior britânicas. Ele é suficientemente flexível para ser adequado a outros tipos de instituição dentro e fora do Reino Unido. O estudo foi financiado pelo JISC.

O modelo está inserido em um relatório que contém sugestões estratégicas acompanhadas de breves descrições dos principais recursos facilitadores, a fim de se implementar uma política de preservação digital. Os autores destacam a importância de relacionar o desenvolvimento de uma política de preservação digital com as diretrizes e estratégias administrativas mais importantes de uma instituição, pois esse tipo de política não produz efeito isoladamente.

O relatório contém duas ferramentas: a) um modelo de *framework* para uma política de preservação digital desenvolvido com base no exame de políticas de preservação digital existentes no Reino Unido e em outros países; b) uma série de mapeamentos das ligações entre a preservação digital e as estratégias em áreas-chave (pesquisa, ensino-aprendizagem, bibliotecas e arquivos) da administração das universidades e faculdades inglesas.

A ferramenta de preservação contém duas seções: a que delineia o modelo de política e a que molda a implementação deste modelo. A primeira está em um nível superior (*high level*); não possui detalhes técnicos; evidencia, porém, pontos fundamentais a serem considerados no princípio de uma política de preservação digital. A segunda inclui orientações técnicas. Deveria compor uma parte significativa de uma política de preservação. Os autores assinalam que o formato de política apresentado é um modelo acompanhado de orientações. Dessa forma, dever-se-á adotar uma abordagem seletiva quando se pretender criar uma política com base no modelo proposto, levando-se em conta as cláusulas das duas seções de acordo com a necessidade e realidade da instituição onde for desenvolvida a política.

O modelo é apresentado de forma sumária no quadro 5, com as cláusulas de política do nível superior, e no quadro 6, com as cláusulas de implementação (nível

inferior). No relatório, os quadros são acompanhados por notas explicativas, estudos de caso e exemplos de cláusulas individuais. As cláusulas são numeradas para facilitar o cruzamento de referências com as notas explicativas.

Quadro 5– Cláusulas do nível superior de uma política de preservação digital

	Cláusula	**Descrição**
5.1	Declaração de princípios	Aponta o modo com que a política de preservação digital pode servir às necessidades da organização e os benefícios que ela trará.
5.2	Ligações contextuais	Destaca a forma com que essa política se integra dentro da organização e com que ela se relaciona com outras estratégias e políticas de alto nível.
5.3	Objetivos da preservação	Informa os objetivos da preservação e o modo de serem eles apoiados.
5.4	Identificação do conteúdo	Caracteriza o que é o escopo geral da política, em termos de conteúdo, e seu relacionamento com os objetivos do desenvolvimento das coleções.
5.5	Responsabilidade nos procedimentos	Identifica as responsabilidades de alto nível para a política e provê o reconhecimento das obrigações mais importantes diante da preservação dos recursos institucionais mais significativos.
5.6	Orientação e implementação	Dita cláusulas de orientação e implementação que estabelecem o modo de implementar a política de preservação e/ou saber onde os procedimentos e orientações adicionais estão disponíveis em documentação à parte ou com o *staff*. As cláusulas e os tópicos importantes do quadro 6 (Implementação) podem ser usados, como se requer, para inserirem-se aqui e/ou proverem um *framework* para uma documentação à parte.
5.7	Glossário	Lista definições, se for necessário.
5.8	Controle de versão	Procede a detalhes históricos e bibliográficos da versão. Adiciona a data da versão, assim como a pretensão de duração e o processo de revisão.

Fonte: adaptado de Beagrie *et al.* (2008, p. 16)

No quadro 5, vale destacar a cláusula 5.2 – ligações contextuais −, pois ela se refere à relação da política de preservação digital com outras políticas e estratégias de alcance mais amplo na instituição. Os autores do modelo consideram que o esforço de implementação de uma política de preservação só vale a pena se ela estiver conectada ao núcleo das diretrizes e estratégias do negócio institucional. Em outras palavras: essa política não pode ser criada de forma isolada. Por essa razão, eles se empenharam no mapeamento de outras estratégias centrais das universidades, incluindo a pesquisa, o ensino-aprendizagem, as bibliotecas e o gerenciamento de registros. Esses mapeamentos

e a metodologia para estudá-los são discutidos de forma breve na segunda seção do relatório.

Os quadros foram construídos com base em uma pesquisa documental que procurou identificar os principais temas nas políticas de preservação digital existentes: os objetivos da preservação, declaração de missão, ligações com outras políticas, aporte financeiro, recursos humanos, questões de propriedade intelectual e aspectos técnicos.

Quadro 6 – Cláusulas do nível inferior de uma política de preservação digital

	Cláusula	**Descrição**
6.1	Responsabilidades financeiras e responsabilidades do *staff*	Esta seção apresenta quem é responsável pela preservação digital dentro da organização. Ela também apresenta a sustentabilidade financeira e o modo com que a política se situa no planejamento financeiro da organização.
6.2	Propriedade intelectual	Esta cláusula mostra o conhecimento sobre as questões de *copyright* e o modo com que a instituição planeja reconhecer e tratar essas questões-chave.
6.3	Serviços distribuídos	Em algumas situações, pode ser mais conveniente ou mais econômico, terceirizar algumas ou todas as atividades de preservação.
6.4	Conformidade com os padrões	Listam-se aqui os padrões adotados pelo repositório.
6.5	Revisão e certificação	Uma descrição da frequência com que a política será revisada. Por exemplo: semestral, anual, bienal etc.
6.6	Auditoria e avaliação de risco	Procedimentos a serem adotados para se cumprirem auditorias padronizadas e reconhecimento dos riscos enfrentados pela política.
6.7	*Stakeholders*	Identificação de todas as partes envolvidas na política e nos procedimentos a serem adotados para a implementação da política.
6.8	Estratégias de preservação	Um guia de estratégias de preservação e um guia de implementações técnicas adotadas.

Fonte: adaptado de Beagrie *et al.* (2008, p. 24-25)

O quadro 6 orienta os procedimentos[50] a serem adotados para a implementação das ações de preservação. Os autores não pretenderam detalhar as estratégias de preservação, mas ajudar os leitores a fixarem conceitos no âmbito mais relacionado com o planejamento da preservação, uma vez que esta é dependente de outras atividades da instituição. Essa questão fica mais clara quando lemos as explicações pormenorizadas da

[50] A diferença entre procedimento e política tem sido explicada de forma esquemática no portal do Governo do Canadá no tópico *What is a Procedure?* da *webpage Concepts for Developing Digital Preservation Policies.* Diponível na WWW: http://canada.pch.gc.ca/eng/1445528228146/1445528228149

última cláusula de implementação denominada de estratégias de preservação. Esta cláusula sugere um dos dois caminhos seguintes: a) realizar uma abordagem pelo ciclo de vida do objeto digital durante as etapas de implementação na seguinte ordem: seleção, conversão, recepção, verificação, determinação das propriedades significantes, ingestão, metadados, armazenagem, técnicas de preservação e acesso; b) optar pela estruturação terminológica do Modelo de Referência OAIS. Se verificarmos o catálogo de políticas do projeto SCAPE (Sierman, Jones e ElstrØm, 2014), essa claúsula poderá ser complementada com o elemento de política *Define Preservation Strategies,* que descreve as estratégias de migração, emulação, preservação de *software/hardware* e filmagem, além de apresentar outros elementos norteadores em seu *template*. A complementaridade entre os elementos de política do SCAPE e o modelo proposto por Beagrie *et al.* (2008) pode ser verificada, ainda, com as demais cláusulas de implementação.

3.3 – O catálogo de elementos de política de preservação do projeto SCAPE

O projeto SCAPE é patrocinado pelo *Seventh Framework Programme for Research and Technological Development* da União Europeia. Foi concluído em setembro de 2014 (King, 2014). Aborda a preservação digital por meio de quatro subprojetos (SCAPE, 2014): *Testbeds; Preservation Components; Platform; Planning and Watch.* Este último é o responsável pela criação do Catálogo de Elementos de Política de Preservação (*Catalogue of Preservation Policy Elements*).

O catálogo é parte do *framework* de política do projeto SCAPE. Este tem como um de seus principais objetivos o de fazer com que as funções de planejamento e monitoramento (*Planning and Watch*) utilizem *workflow*s compatíveis com uma política automatizada. O *framework* é estruturado em três níveis (figura 7): o nível de política superior dentro de uma instituição (*Guindance Policy*); o nível onde as políticas são definidas de modo mais específico (*Preservation Procedure Policies*); o nível onde são criadas declarações que servirão de base para um *workflow* automatizado (*Control Policies*). O catálogo descreve o nível intermediário (*Preservation Procedure Policies*) detalhadamente e com referências aos outros níveis. A conexão desses três níveis torna mais fácil o processo de criar políticas, uma vez que ele ajuda a conscientização sobre a

necessidade de formulá-las com um maior grau de detalhamento, assim como à preparação dessas políticas para se tornarem legíveis a máquinas (Sierman, Jones e ElstrØm, 2014).

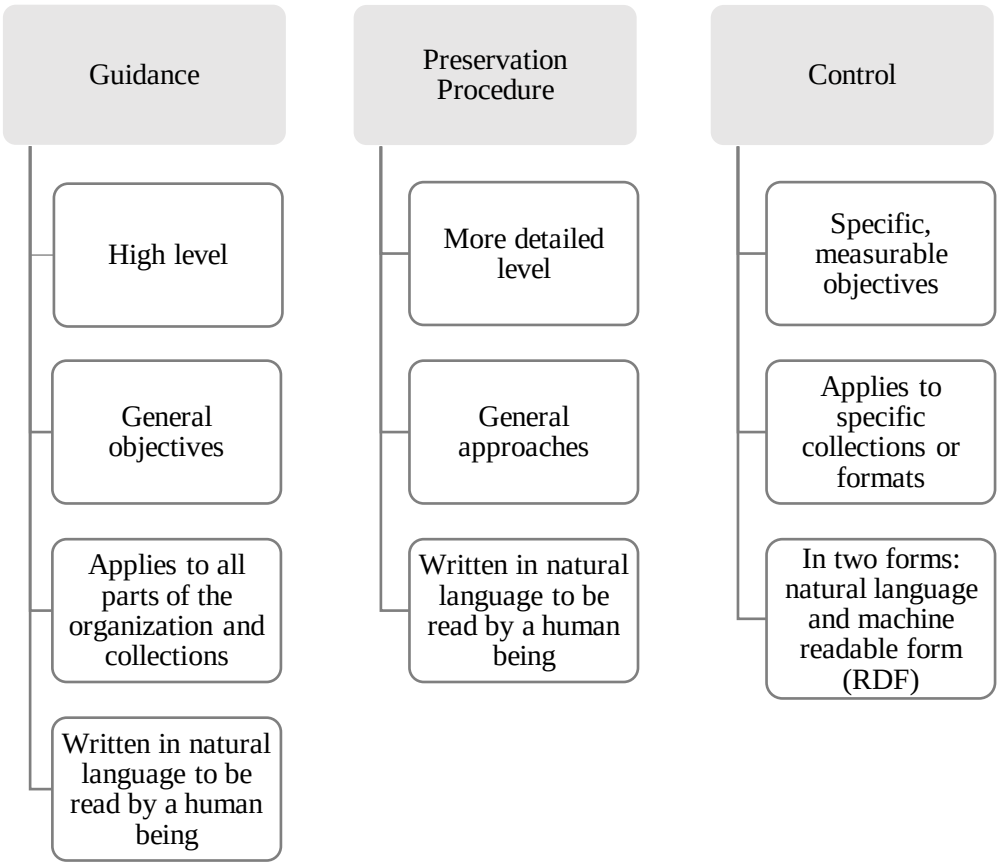

Figura 7 – Níveis de política de preservação digital identificados no SCAPE

Fonte: Sierman; Jones & ElstrØm (2014, p. 7)

Os níveis de política são descritos da seguinte forma:

> **Políticas de orientação** − Neste nível, a organização descreve os objetivos gerais da preservação da sua coleção (ou coleções) a longo prazo. Um exemplo é aquele em que uma organização determine que a infraestrutura estabelecida para suportar a preservação digital seja guiada pelo modelo OAIS.
> **Políticas no nível de procedimentos de preservação** − Estas políticas descrevem a abordagem que a organização adotará para alcançar as metas estabelecidas no nível mais alto. Elas serão detalhadas o suficiente para ser inseridas em *design* de processos e *workflows*, mas podem ser, ao mesmo tempo, de interesse para as coleções de um modo geral. Esse é o nível que tematiza o Catálogo de Elementos de Política.
> **Políticas de controle** − Neste nível, as políticas formulam os requisitos para uma coleção específica, para uma ação de preservação específica ou para uma comunidade de usuários específica. Este nível pode ser legível aos humanos, mas também deveria ser a máquinas e estar pronto para uso. Assim, ele pode ser usado em ferramentas de planejamento e monitoramento automáticos para

> assegurar que as ações e os *workflows* de preservação escolhidos descubram os requisitos específicos identificados para uma dada coleção digital. (Sierman; Jones & ElstrØm, 2014, p. 7, tradução nossa).

Esses autores sugerem que as políticas de orientação (*guidance policies*) sejam disponibilizadas para a comunidade de usuários, enquanto afirmam que as políticas de controle (*Control policies*), por serem muito minuciosas e possivelmente possuírem informações operacionais, devem ser de acesso restrito. A disponibilidade das políticas no nível de procedimentos de preservação (*preservation procedures*) depende da organização e da sua comunidade de usuários.

Os elementos das políticas de orientação recebem a seguinte denominação:

- Autenticidade: refere-se às medidas para se estabelecer a autenticidade.
- Preservação de *bit*: mensuração da integridade de *bit*, locais diferentes para armazenamento, regularidade de verificação, custos, níveis de segurança etc.
- Preservação funcional: abrange diversas estratégias com o objetivo de preservar diferenciados tipos de material digital.
- Objeto digital: especificidades, originalidade, propriedades significantes etc.
- Metadados: políticas relacionadas com os metadados.
- Acesso: explica-se como o objeto será visto e apontam-se métodos para reuso do material digital e *harvesting* etc.
- Direitos: políticas relacionadas com o acesso, preservação etc.
- Padrões: o emprego de padrões.
- Organização: políticas relacionadas com o comportamento e as tarefas da organização de acervos.
- Auditoria e certificação: *Data Seal of Approval*, DIN 31644, ISO 16363.

Cada política no nível de orientação é acompanhada de um referencial teórico e vários elementos de procedimento de preservação (*preservation procedure policy*). Assim, por exemplo, o elemento de política denominado de autenticidade consiste nos seguintes procedimentos de preservação: a integridade, a confiabilidade e a proveniência. Cada política de procedimento possui um *template* com 10 campos explicativos e tem por finalidades:

- Identificar os *stakeholders* envolvidos com a política.

- Identificar as etapas do ciclo de vida do material digital, para os quais a política é relevante.
- Relacionar um dado elemento de política com o seu nível mais alto e ser consistente.
- Identificar outros níveis de política de orientação que podem influir em um elemento no nível de procedimentos de preservação.
- Estimular o debate dentro da instituição com questões sugeridas pelo próprio *template*.

O catálogo é fundamental para a implementação das políticas de preservação digital, especialmente aquelas voltadas para a preservação em larga escala. Mesmo nos casos que não se pretender automatizá-las, ele é um imprescindível complemento do nível mais alto, devido ao seu grau de detalhamento técnico, como pudemos ver na análise do estudo de Beagrie *et al.* (2008).

3.4 – O modelo de política do InterPARES

O projeto InterPARES desenvolveu um conjunto de recursos educacionais em preservação para profissionais que trabalham com registros digitais. Um de seus objetivos é o de suprir os currículos universitários com conteúdo e estrutura para cursos voltados à gestão e preservação de registros digitais.

A política do InterPARES, conforme é definida, deverá ser chancelada pelo nível mais alto da hierarquia administrativa, estar em consonância com a gestão e com a cultura da instituição e servir de apoio à implementação de um programa de preservação digital. Neste sentido, a primeira parte do modelo de *framework* do JISC (quadro 5) é coerente com essa definição. A cláusula 5.1, por exemplo, deve estar alinhada com as declarações de missão da organização e prover um ponto de partida para o restante da política, conforme explicações adicionais sobre esta cláusula, denominada de declaração de princípios.

O modelo de política do InterPARES foi desenvolvido em parceria com o *International Council on Archives*, numa iniciativa educacional denominada *Digital*

Records Pathways: Topics in Digital Preservation, que contém uma série de 8 módulos. O módulo 2, intitulado *Developing Policy and Procedures for Digital Preservation*, assevera vários aprendizados. Dentre estes, destacamos: o de capacitar o estudante a entender os propósitos e benefícios de uma política de preservação digital e os procedimentos conexos; o de aprender a distinguir o sentido de política e o de procedimentos. Neste caso, o módulo ensina: "(...) onde a política articula conceitos de alto nível que delineiam e guiam a responsabilização, os procedimentos dão aos criadores de registros atividades e responsabilidades concretas para assegurar a criação de registros autênticos, confiáveis, acurados e usáveis." (InterPARES, 2012, p. 33, tradução nossa).

Acerca dos propósitos de uma política de preservação digital, o módulo 2 considera que, apesar de vários projetos internacionais estarem pesquisando soluções para a preservação digital e apesar do desenvolvimento de ferramentas tecnológicas para este fim, a tecnologia é apenas uma parte da solução. A preservação digital, para surtir efeito, deveria apoiar os objetivos e as metas de uma organização através de *frameworks* e políticas institucionais.

Uma política de preservação digital assegura o acesso contínuo aos registros digitais, de modo que facilita as tomadas de decisão em uma instituição. As pesquisas demonstram que não é possível preservar o objeto digital, mas apenas sua reprodutibilidade. A autenticidade, confiabilidade e acurácia de um objeto depende de um *framework* apropriado. A capacidade de manter sua autenticidade ao longo do tempo deve ser tratada quando tal objeto estiver sendo criado. O projeto InterPARES 2 criou um *framework* de princípios para ajudar na criação de políticas, estratégias e padrões que são plenamente adaptáveis em qualquer país, equilibrando a perspectiva cultural, social e jurídica, além de ser robusto o suficiente para fundamentar qualquer documento de política dessa natureza.

O módulo 2 recomenda situar os princípios do InterPARES 2, sob a perspectiva de um objetivo de política abrangente que estabeleça a ligação entre os registros digitais e o negócio da instituição. Tais princípios estão implícitos no *template* de política descrito a seguir, de forma resumida (InterPARES, 2012):

a) Princípios

Uma política de preservação digital determina os princípios gerais que orientam a implementação de um programa de preservação e gestão de objetos digitais, garantindo a sua confiabilidade, autenticidade e acessibilidade a longo prazo. Ela norteia a gestão desses objetos durante o período que exceder a vida útil da tecnologia que criou os objetos; atribui responsabilidades a quem cria e usa os objetos digitais e deve ser redigida em uma linguagem clara e concisa. Quando forem empregados termos eminentemente técnicos, os leitores deverão ser remetidos a um glossário. Finalmente, uma política desse tipo deve estar sujeita a reavaliações periódicas e contínuas.

b) Elementos da política

A segunda parte do modelo é constituída por nove elementos de política: propósitos e objetivos; escopo; mandato; declaração da política; papéis e responsabilidades; definições; fontes relacionadas; controle de versão; revisão da política.

- Propósitos e objetivos

Os objetivos de uma política devem constar na seção introdutória e ser alinhados com os objetivos da instituição. As políticas que governam a criação, manutenção e preservação dos registros digitais devem posicionar-se em relação às questões de confiabilidade, acurácia e autenticidade.

- Escopo

O escopo de uma política deve indicar os objetos digitais que são protegidos por ela e indicar os indivíduos e departamentos para quem a política se destina.

- Mandato

O mandato de qualquer setor de uma organização emitente da política deverá ser especificado. A inclusão de um mandato indicará que o conselho administrativo emissor da política tem a autoridade para determiná-lo. A política supre as necessidades de negócio da organização.

- Declaração da política

A declaração política faz o enquadramento das responsabilidades dos criadores dos objetos digitais e assegura que os objetos são produzidos com acurácia e mantidos com autenticidade. Ela deve ser fundamentada nas necessidades do negócio da instituição e não na solução tecnológica que visa a suprir tais necessidades. A política será revisada periodicamente e modificada à medida que a instituição for evoluindo.

- Papéis e responsabilidades

Esta seção determina a responsabilidade pela implementação da política no âmbito mais abrangente da estrutura administrativa, identifica as partes interessadas (*stakeholders*) e lhes atribui as responsabilidades permanentes para estas assegurarem o compromisso com a política em todos os níveis da instituição. É nesta seção que o conjunto de responsabilidades é examinado e definido, de modo que diferencia a responsabilidade daqueles que exercem as ações de preservação dos objetos digitais e a daqueles que os chefiam.

- Definições

Esta seção deve conter um glossário de termos específicos de um domínio de conhecimento ou um glossário da própria instituição, usados na política, especialmente se esses termos diferem da linguagem do dia a dia da instituição.

- Fontes relacionadas

Uma política deve observar a legislação local ou nacional e seguir os padrões mais usuais e as práticas mais difundidas. A legislação, as melhores práticas e padrões devem ser referenciados na política.

- Controle de versão

Cada política deve ter uma informação sobre o controle de versão para assegurar que os interessados estão pautando-se pela versão mais atual. Cumpre fornecer algumas informações essenciais:

  - O número da versão.
  - A data inicial para a política ter efeito.
  - Indicar a data, se a política estiver desatualizada.
  - Cada política deve indicar a versão imediatamente anterior que ela atualizou e a versão antiga deve indicar a versão atualizada.

- Revisão da política

As políticas devem ser aprovadas no nível da gestão mais alto da hierarquia institucional que reconhece a importância do tema. Como os materiais digitais testemunham as atividades de uma organização e a tornam responsável pelas ações sobre eles, a esfera mais alta da administração deveria aprovar uma política de preservação digital. Uma revisão da política poderia ser realizada por uma assessoria jurídica, para

esta garantir que tal política está amparada na legislação pertinente e se harmoniza com as políticas organizacionais de gestão de registros, de acesso à informação e de privacidade.

As políticas devem ser revisadas periodicamente, para se garantir que elas continuam provendo a melhor orientação, com vistas a atingir as metas institucionais. Uma política deve verificar como, quando e por quem ela será revisada.

Esta seção deve conter as seguintes informações:

- Os indivíduos e departamentos responsáveis pela aprovação da política;
- O período de tempo entre as revisões;
- A data da última revisão;
- A data da aprovação da política;
- A data da próxima revisão.

Esse *template* é precedido pela apresentação de uma metodologia de pesquisa-ação, cujos fundamentos assentam em uma aplicação interativa de práticas que vão desde a reunião de dados e o diálogo colaborativo até às tomadas de decisão em grupo. A metodologia é acompanhada de exercícios. O capítulo 5 do módulo 2 apresenta um estudo de caso e o capítulo 7 é dedicado ao exercício de análise de políticas. O que se depreende desse módulo é a sua importância em ser um suporte didático para o aprendizado da elaboração de uma política. Ele se diferencia dos demais modelos apresentados no presente trabalho, porque, além de propor um *template* de política, disponibiliza um *workflow* que funciona como uma ferramenta para o refinamento de uma política que estiver sendo elaborada. Na verdade, uma equipe iniciante deveria estudar esse módulo ao mesmo tempo que se constrói uma política com base em outros modelos, como o de Beagrie *et al.* (2008) e o Catálogo de Elementos de Política de Preservação do projeto SCAPE.

Capítulo 4 – Metodologia

O método em pesquisa é a escolha de procedimentos sistemáticos feita para se descreverem e se explicarem fenômenos. Tais procedimentos se avizinham do método científico, que compreende a delimitação de um problema e a realização de observações e de suas interpretações, de acordo com as relações encontradas, com base nas teorias disponíveis, quando estas forem possíveis. Assim, destacam-se dois grandes métodos: o quantitativo e o qualitativo. Diferenciam-se pela sistematização característica de cada um deles e, mais ainda, pelo modo de abordar um problema.

O método quantitativo consiste na quantificação nas fases de coleta e tratamento de informações, utilizando-se técnicas estatísticas básicas (percentual, média, desvio-padrão etc.) ou complexas (coeficiente de correlação, análise de regressão etc.). Já o método qualitativo não utiliza as técnicas estatísticas para fundamentar a análise de um problema. Esse método é uma opção do investigador na abordagem de um problema e propicia o entendimento da natureza de um fenômeno social (Richardson, 2015).

Os alcances da pesquisa decorrem da revisão da literatura e da perspectiva de estudo e, para combinarem os elementos no referido estudo, dependem dos objetivos pretendidos. Dois tipos de alcance nos interessaram particularmente: os estudos exploratórios e os estudos descritivos:

> Os **estudos exploratórios** são realizados quando o objetivo é examinar um tema ou um problema de pesquisa pouco estudado, sobre o qual temos muitas dúvidas ou que não foi abordado antes. Ou seja, quando a revisão da literatura revelou que existem apenas orientações não pesquisadas e ideias vagamente relacionadas com o problema de estudo ou, ainda, se queremos pesquisar sobre temas e áreas a partir de novas perspectivas. […]
> Os **estudos descritivos** buscam especificar as propriedades, as características e os perfis de pessoas, grupos, comunidades, processos, objetos ou qualquer outro fenômeno que se submeta a uma análise. Ou seja, pretendem unicamente medir ou coletar informação de maneira independente ou conjunta sobre os conceitos ou as variáveis a que se referem, isto é, seu objetivo não é indicar como estas se relacionam. […]. (Sampiere, Collado, Lucio, 2013, p. 101-102).

Os estudos recentes sobre repositórios institucionais no Brasil não revelaram a existência de políticas de preservação digital, especialmente nas universidades federais. Tal lacuna motivou a sistematização da nossa pesquisa. Esta consistiu em percorrermos os objetivos específicos até levantarmos as características da cultura de preservação digital nesse tipo de instituição de ensino superior. Por conseguinte, adotando o método

qualitativo, iniciamos a pesquisa de modo exploratório e analisamos os dados coletados de forma descritiva, com o apoio da técnica de estatística básica. Nesse contexto, os objetivos específicos são caracterizados da seguinte forma:

- A análise das grandes mudanças na academia com a emergência da ciberciência se orienta por um método qualitativo de alcance exploratório;
- A caracterização da origem e da função dos RIs segue um método qualitativo de alcance descritivo;
- A análise dos modelos de *framework* de política de preservação digital, elaborados por iniciativas norte-americanas e europeias, é de natureza qualitativa com alcance exploratório;
- A verificação da existência de uma política de preservação digital para os RIs das universidades federais brasileiras é realizada por meio do método quanti-qualitativo com um alcance descritivo;
- A apuração da percepção dos administradores dos RIs das universidades federais sobre as questões de preservação digital segue o método quantiqualitativo de alcance descritivo.

4.1 – O universo da pesquisa

Coutinho (2015, p. 90) salienta:

> Nem sempre é necessário o investigador constituir uma amostra para o seu estudo, caso, por exemplo, da investigação histórica ou da investigação-ação que partem sempre de um grupo específico para a análise, ou mesmo de estudos em que o grupo-alvo (*target group*) coincide a população, ou seja, é analisado na sua totalidade.

O Brasil possui 63 universidades federais. O universo (ou população) da nossa pesquisa é constituído por todas as 38 universidades federais que possuem um repositório institucional. Dentre elas, 26 cadastraram seus respectivos repositórios no OpenDOAR. Os demais foram descobertos explorando-se os portais de todas as outras universidades. Nos casos em que não foi possível encontrar um RI, fizemos contato com a biblioteca central por *e-mail* e por chamada telefônica, a fim de confirmar a existência ou inexistência de um repositório.

4.2 – A coleta e a análise dos dados

Sampieri, Collado e Lucio (2013) ensinam que a coleta de dados pelo método qualitativo ocorre nos ambientes naturais e de rotina dos participantes ou nas unidades de análise. O instrumento de coleta de dados é o pesquisador, pois ele, além de fazer a análise, é o próprio meio de obtenção de informação. Para tanto, o pesquisador utiliza várias fontes de dados: entrevistas, questionários, observações diretas, documentos, material audiovisual dentre outras. Os autores listam onze tipos de unidade de análise: significados, práticas, episódios, encontros, papéis, relações, grupos, organizações, comunidades, subculturas e estilos de vida. No interesse do presente trabalho, destacamos as unidades de análise referentes às práticas e às organizações:

a) As práticas constituem uma unidade de análise comportamental relacionada com uma atividade contínua, identificada pelos membros de um sistema social como uma rotina.
b) As organizações são unidades criadas com um fim coletivo. A análise delas, foca-se muitas vezes na origem, no controle, nas hierarquias e na cultura.

Nesta pesquisa, entendemos as universidades federais brasileiras como organizações sociais[51] que possuem uma cultura de comunicação e informação científica ainda em fase incipiente, no que diz respeito ao movimento de acesso aberto. Os RIs são repositórios digitais de acesso aberto e as universidades precisam planejar a manutenção e a preservação, a longo prazo, do acervo dos seus repositórios. Procurar apurar as práticas de preservação adotadas por profissionais responsáveis por um RI é essencial para entendermos o que está moldando a preservação digital nesse tipo de sistema nas universidades federais. Por essa razão, além da coleta de dados disponíveis no OpenDOAR e nos *websites* dos RIs, submetemos um questionário *online* aos administradores dos repositórios.

Coutinho (2015) explica que a notação é um processo em que o investigador observa comportamentos, fenômenos ou documentos e registra as ocorrências ou lista as

[51] Para uma melhor compreensão sobre a universidade tida como organização social em oposição a uma instituição social, ver Chauí (2003).

características desses elementos. Esse processo parte da observação, pode ser utilizado em qualquer tipo de pesquisa e é central nos estudos descritivos.

Nossas observações partiram das consultas realizadas no OpenDOAR para verificarmos se haveria algum dado relativo a uma política de preservação digital nos 26 repositórios cadastrados. Após essa verificação, consultamos as *homepages* de todos os 38 repositórios, nas quais encontramos o principal documento de análise: a Política de Informação Institucional (PII). Por outro lado, não foi possível encontrar documentos que estabelecessem alguma atividade de preservação digital. Assim, elaboramos um questionário *online* com perguntas fechadas e abertas, a fim de conhecer a cultura de preservação digital nas universidades federais. O questionário *online* justifica-se, porque as universidades federais estão localizadas em todos os estados da federação. Essa fonte de coleta de dados tornou viável a pesquisa, pois minimizou-lhe o custo operacional e financeiro.

Nas perguntas fechadas, as opções de respostas são previamente delimitadas, o que facilita a codificação e a análise. Elas podem ser dicotômicas ou incluir várias opções de respostas. De modo contrário, as perguntas abertas não delimitam tais opções. Ajudam-nas a suprir a falta de informações sobre as possíveis respostas dos inqueridos (Sampieri, Collado e Lucio, 2013). O questionário aplicado nesta pesquisa possui perguntas fechadas dicotômicas, perguntas com opções de respostas de múltipla escolha e de repostas múltiplas, assim como perguntas abertas. Ele foi estruturado em duas partes, com base no modelo de *framework* do projeto JISC (Beagrie *et al.*, 2008). Possui vinte e quatro questões. A primeira parte relaciona as questões que buscam levantar os elementos do nível superior de uma política de preservação digital; a segunda procura os elementos do nível inferior (ou de implementação) desse tipo de política.

Realizamos dois pré-testes com a cooperação de uma universidade respondente e pioneira na implantação de repositórios institucionais. O questionário (vide anexo) foi construído com a ferramenta de formulários do Google e enviado aos trinta e oito administradores dos RIs. Dessa forma, obtivemos as oito primeiras respostas espontâneas em cerca de uma semana. A coleta durou trinta e seis dias. Teve início no dia 10 de março de 2016 e terminou em 14 de abril de 2016. Para chegarmos à taxa de 100% de respostas, procuramos os administradores dos RIs por meio de chamadas telefônicas, a partir do dia 21 de março de 2016.

Coutinho (2015, p. 106, sublinhado da autora) explica:

> Embora o termo **análise** possa ser usado na linguagem de investigação social com outros significados, como procedimento ligado à recolha de dados numa investigação, refere o processo em que o investigador, mais do que observar (caso da descrição) procura sim "... inferir traços, processos, significados e relações" (Charles, 1998, p. 154). Ou seja, o investigador não vai para o terreno observar tudo o que passa, mas apenas aquilo que interessa no contexto do seu estudo, ou seja, os objetivos específicos que procura alcançar: (...).

A análise empreendida nesta pesquisa procurou, primeiramente, verificar se alguma política de preservação digital fora derivada das PIIs encontradas nas *homepages* dos repositórios. Em seguida, após a aplicação do questionário, analisamos os sentidos dos dados colhidos e quantificados; relacionamos as descobertas com o referencial teórico abordado. No terreno desta pesquisa, poderíamos observar outros aspectos, tais como: a formação dos administradores dos repositórios, seus respectivos perfis sociais etc. Porém, o que nos interessou foi apenas conhecer o contexto da cultura de preservação digital nas universidades federais.

Capítulo 5 – Os repositórios institucionais das universidades federais brasileiras: análise e discussão dos resultados

5.1 – Diretory of Open Access Repositories (OpenDOAR)

Em setembro de 2005, o IBICT publicou o *Manifesto Brasileiro de Apoio ao Acesso Livre à Informação Científica*. Dentre as suas declarações, o documento aconselhava as instituições acadêmicas brasileiras a se comprometerem com a criação de repositórios temáticos e institucionais, de acordo com o paradigma do acesso aberto (IBICT, 2005). Consequentemente, o próprio IBICT desenvolveu uma iniciativa no nível da legalidade e outra no nível da infraestrutura. No nível da legalidade, o IBICT participou da elaboração de um projeto de lei no Senado Federal, sob o número 387/2011 (Rollemberg, 2011). Tal projeto visava a obrigar as instituições públicas de ensino superior e as de pesquisa a criarem seus repositórios institucionais. No nível da infraestrutura, a iniciativa consistiu na doação de dezenas de computadores para as universidades com a função de repositório digital. Neste sentido, a contrapartida das instituições de ensino superior seria a configuração do repositório e a adoção de um modelo de política denominado de Política de Informação Institucional (PII), cujo propósito é o de povoar o repositório e preservar a produção científica (Kuramoto, 2009).

Existem 63 universidades federais.[52] Dentre elas, 38 (60%) possuem repositórios. Dentre estes, 26 estão cadastrados no OpenDOAR, mas nenhum deles registrou qualquer aspecto relacionado com uma política de preservação digital na ferramenta de políticas daquele diretório. Apenas 16 universidades criaram uma PII e a publicaram no *website* dos seus repositórios.

A listagem a seguir é um levantamento do que encontramos sobre a política de preservação digital dos repositórios cadastrados no OpenDOAR e nas PIIs disponíveis em seus respectivos *websites*.

[52] No dia 9 de maio de 2016, o governo assinou um projeto de lei para a criação de 5 universidades federais: Universidade Federal de Catalão, Universidade Federal de Jataí, Universidade Federal do Delta do Parnaíba, Universidade Federal do Norte de Tocantins e Universidade Federal de Rodonópolis.

- Universidade Tecnológica Federal do Paraná (UTFPR) — A última revisão de suas políticas no OpenDOAR ocorreu em outubro de 2013 e o *status* da política de preservação está definido como "*explicitly undefined*". Uma PII foi publicada em dezembro de 2009 no *website* do repositório (UTFPR, 2009). A política pretende facilitar a preservação e considera a formulação de uma política específica para preservar o material armazenado, mas não apresenta qualquer guia para orientar o planejamento de tal preservação.
- Universidade Federal da Bahia (UFBA) — A última revisão de suas políticas no OpenDOAR ocorreu em janeiro de 2012. Seu *status* também foi classificado como "*explicitly undefined*". A PII do repositório manifesta uma intenção geral em preservar a produção científica da universidade (UFBA, 2010); entretanto, não especifica os meios para pôr tal preservação em prática.
- Universidade Federal da Grande Dourados (UFGD) — A última revisão do repositório no OpenDOAR ocorreu em agosto de 2012. Semelhante aos repositórios já mencionados, o *status* de sua política de preservação foi estabelecido como "*explicitly undefined*". Uma PII foi publicada, apesar de não ter sido formalizada como um documento oficial (UFGD, 2010). Vale dizer que a preservação da produção tecnocientífica também é uma preocupação manifestada nesse documento.
- Universidade Federal de Pelotas (UFPEL) — A última revisão das políticas no OpenDOAR foi realizada em maio de 2013. A existência de sua política de preservação ainda não havia sido confirmada. A PII da universidade foi promulgada em junho de 2010 e definiu o RI como um sistema de informação para armazenar, preservar, organizar e difundir a produção acadêmica da instituição. Ademais, os artigos 7 e 8 dessa política estabeleceram responsabilidades e prerrogativas para o *staff* administrativo. Dessa forma, esse grupo deve lidar com os metadados, formatos, migração e preservação do conteúdo, com cautela (UFPEL, 2010). Entretanto, nenhuma outra política relacionada com o desenvolvimento de um programa de preservação foi encontrada no âmbito deste estudo.

- Universidade Federal do Rio Grande do Sul (UFRGS) — A última revisão de suas políticas no OpenDOAR foi realizada em julho de 2010 e, 5 anos depois, a descrição da preservação ainda era dada como "*explicitly undefined*". Não obstante, essa universidade publicou uma PII, por meio da qual, manifesta uma preocupação com a preservação digital da memória institucional no longo prazo (UFRGS, 2010). Apesar dessa preocupação, não há indicação do modo com que a universidade poderia implementar as atividades de preservação. Cabe observar que essa PII delineia uma seção para definir os requisitos de metadado, que é um elemento fundamental para a preservação digital; no entanto, a definição da tipologia dos metadados foi atribuída aos *stakeholders*. Mesmo assim, nenhum documento específico foi referenciado para orientar os administradores e usuários do repositório.
- Universidade Federal do Rio Grande do Norte (UFRN) — A última revisão das políticas do repositório no OpenDOAR, realizada em março de 2011, registrou uma situação semelhante à dos repositórios já mencionados, isto é, o *status* da política de preservação desse repositório também foi registrado como "*explicitly undefined*". A PII da universidade destaca a necessidade de preservar a produção tecnocientífica da instituição, mas não explica como uma iniciativa dessa natureza poderia ser levada a efeito (UFRN, 2010).
- Universidade Federal de Sergipe (UFS) — A última revisão de suas políticas no OpenDOAR data de janeiro de 2013 e o *status* da política de preservação foi classificado como "*explicitly undefined*". O artigo 5º da Política de Acesso Livre da Produção Científica da UFS estabelece que os autores devem autorizar a prerservação habilitada pelo repositório. Contudo, nada é dito sobre o modo com que tal preservação poderia ser levada a cabo (UFS, 2010).
- Universidade Federal dos Vales do Jequitinhonha e Mucuri (UFVJM) — A última revisão no OpenDOAR ocorreu em maio de 2011 e o *status* da política de preservação está classificado como "*explicitly undefined*". A Política de Funcionamento do Repositório Institucional dessa universidade foi regulamentada em 2010. Ele é semelhante às iniciativas definidas por outras universidades como uma PII, apesar de ter uma denominação diferente. Ela também considera a preservação como um de seus objetivos, conforme o que é

estabelecido no artigo 2º, mas não especifica como essa preservação poderia ocorrer (UFVJM, 2010).

- Universidade Federal do Rio Grande (FURG) — A última revisão do OpenDOAR decorreu em junho de 2011. Como foi observado nos casos anteriores, a descrição da preservação é estabelecida como "*explicitly undefined*". Sua PII leva em consideração a necessidade de preservar a produção intelectual da instituição e atribui ao comitê gestor a manutenção de um conjunto de dados atualizados e organizados, a fim de garantir a preservação digital do acervo do repositório (FURG, 2011). Entretanto, nada mais foi acrescentado em termos de preservação digital, especialmente o que diz respeito a uma política para esta finalidade.
- Universidade Federal do Ceará (UFC) — Recebeu sua mais recente revisão no OpenDOAR em setembro de 2014. Sua política de preservação foi classificada como "*explicitly undefined*". A PII da universidade considera a necessidade de preservar a produção tecnocientífica. No artigo 3º há um parágrafo único estabelecendo que os autores devem permitir a instituição preservar seus trabalhos acadêmicos, de acordo com as condições definidas pelo termo de autorização. Um modelo desse termo está disponível para *download* na *homepage* do repositório (UFC, 2011). Neste caso, também é importante assinalar que a referida política não especifica como sua pretensão de preservação seria alcançada.
- Universidade Federal de Lavras (UFLA) — A última revisão no OpenDOAR de suas políticas foi em janeiro de 2013. O *status* de sua preservação apresenta-se como "*not found*". Uma PII foi publicada no *website* do repositório. Estabelece a intenção de preservar a produção acadêmica da universidade (UFLA, 2012). Todavia, ela não indica qualquer política específica que possa efetivar um programa de preservação digital.
- Universidade Federal de Ouro Preto (UFOP) — A última revisão de suas políticas no OpenDOAR ocorreu em março de 2015. Registrou-se o *status* da sua política de preservação como "*explicitly undefined*". A PII do repositório reconhece a necessidade de preservação do acervo, mas no *website* do repositório não há qualquer documento que direcione um programa de preservação digital (UFOP, 2013).

- Universidade Federal do Recôncavo da Bahia (UFRB) — A última revisão de suas políticas no OpenDOAR aconteceu em maio de 2011. A política de preservação recebeu o *status* definido como "*unknown*". A PII da universidade considera a necessidade de preservar a produção tecnocientífica da instituição e no artigo 12º prevê a criação de normas complementares; no entanto, a política não declara qual seria a iniciativa a ser tomada para orientar os *stakeholders* nas atividades de preservação (UFRB, 2013).
- Universidade de Brasília (UNB) — A última revisão das políticas do repositório no OpenDOAR deu-se em janeiro de 2013. O *status* da política de preservação também foi estabelecido como "*explicitly undefined*". Uma PII foi publicada no *website* do repositório. Declara como primeira necessidade a preservação da produção científica da universidade (UNB, 2013). Por outro lado, ela não apontou o caminho que se poderia seguir para se implementar a evocada preservação da produção científica da universidade.
- Universidade Federal de Minas Gerais (UFMG) — Há dois registros do repositório no OpenDOAR: ID 2907 e ID 3457. A última revisão do primeiro registro ocorreu em novembro de 2013 e a do segundo, em agosto de 2015. Nos dois registros, o *status* das políticas de preservação é designado como "*not found*". A *homepage* do repositório contém uma seção denominada "política do repositório". Apesar de essa política prever a preservação dos materiais relevantes e de todos os tipos de formato digital, nenhuma política de preservação específica foi indicada para orientar um programa de preservação digital (UFMG, 2016).

As 11 universidades listadas abaixo não possuem uma PII. Assim como as demais instituições descritas acima, não registraram uma política de preservação no OpenDOAR, nem tampouco publicaram uma política desse tipo em seus respectivos repositórios:

- Universidade Federal do Espírito Santo (UFES, 2016).
- Universidade Federal Fluminense (UFF, 2016).
- Universidade Federal de Goiás (UFG, 2016).
- Universidade Federal do Maranhão (UFMA, 2016).

- Universidade Federal do Mato Grosso do Sul[53] (UFMS, 2016).
- Universidade Federal do Pará (UFPA, 2016).
- Universidade Federal da Paraíba (UFPB, 2016).
- Universidade Federal do Paraná (UFPR, 2016).
- Universidade Federal de Pernambuco (UFPE, 2016).
- Universidade Federal de Santa Catarina (UFSC, 2016).
- Universidade Federal de Uberlândia (UFU, 2016).

Outras 12 universidades possuem RIs, mas estes não estão cadastrados no OpenDOAR. Apenas 2 delas possuem uma PII:

- Universidade Federal da Integração Latino-Americana (UNILA, 2013).
- Universidade Federal de Alagoas (UFAL, 2016).
- Universidade Federal de Itajubá (UNIFEI, 2016).
- Universidade Federal de Juíz de Fora (UFJF, 2016).
- Universidade Federal de Rondônia (UNIR, 2016).
- Universidade Federal de Santa Maria (UFSM, 2016).
- Universidade Federal de São Paulo (UNIFESP, 2016).
- Universidade Federal de Viçosa (UFV, 2016).
- Universidade Federal do Mato Grosso (UFMT, 2016).
- Universidade Federal do Pampa — No artigo 1º de sua PII, a universidade reconhece que é necessário preservar a produção científica da universidade e no artigo 5º determina ao comitê gestor do repositório "manter o conjunto de dados atualizados e organizados, servindo como garantia da preservação digital" (UNIPAMPA, 2015). Todavia, um tal tipo de garantia só poderá ser consolidado por meio de ações mais amplas e direcionadas por compromissos estabelecidos em uma política de preservação digital. Ela nem sequer foi mencionada na PII.
- Universidade Federal do Piauí (UFPI, 2016).

[53] Uma norma para o funcionamento do repositório foi estabelecida por meio de uma resolução *ad referendum* em 2011, mas nunca foi regulamentada pelo Conselho Universitário. A norma está disponível em WWW: http://biblioteca.sites.ufms.br/files/2016/02/dspace_resolucao.pdf.

- Universidade Federal do Tocantins — Sua PII acrescenta normas para o depósito de teses e dissertações e inclui esses tipos de trabalho nas suas pretensões de preservar "(...) a produção intelectual acadêmica, científica e tecnológica institucional em suporte digital." (UFT, 2011). Assim como nas demais PIIs não há referência a uma política de preservação digital.

O quadro 7 apresenta um resumo dos principais dados demonstrados nos parágrafos anteriores, isto é, menciona as universidades que cadastraram seus RIs no OpenDOAR, publicaram uma PII e possuem uma política de preservação digital (PPD). Os campos assinalados com a letra 'x' indicam a ocorrência do dado da coluna correspondente e o símbolo '−' indica a ausência do dado.

Quadro 7 – RIs que possuem um cadastro no OpenDOAR, uma PII e uma PPD

Universidade	Repositório cadastrado no OpenDOAR	Política de Informação Institucional	Política de Preservação Digital
1. UTFPR	X	X	–
2. UFBA	X	X	–
3. UFGD	X	X	–
4. UFPEL	X	X	–
5. UFRGS	X	X	–
6. UFRN	X	X	–
7. UFS	X	X	–
8. UFVJM	X	X	–
9. FURG	X	X	–
10. UFC	X	X	–
11. UFLA	X	X	–
12. UFOP	X	X	–
13. UFRB	X	X	–
14. UNB	X	X	–
15. UFMG	X	–	–
16. UFES	X	–	–
17. UFF	X	–	–
18. UFG	X	–	–
19. UFMA	X	–	–
20. UFMS	X	–	–
21. UFPA	X	–	–
22. UFPB	X	–	–
23. UFPR	X	–	–
24. UFPE	X	–	–
25. UFSC	X	–	–
26. UFU	X	–	–
27. UNILA	–	–	–
28. UFAL	–	–	–
29. UNIFEI	–	–	–
30. UFJF	–	–	–
31. UNIR	–	–	–
32. UFSM	–	–	–
33. UNIFESP	–	–	–
34. UFV	–	–	–
35. UFMT	–	–	–
36. UNIPAMPA	–	X	–
37. UFPI	–	–	–
38. UFT	–	X	–

Fonte: dados da pesquisa, 2017.

Como já foi dito, os repositórios das universidades federais registrados no OpenDOAR não têm uma política de preservação digital, embora todos os que publicaram uma PII declarem a intenção de preservar os materiais armazenados. Dos trinta e oito repositórios analisados, apenas dezesseis (42%) publicaram uma PII —

conforme o modelo proposto pelo IBICT —, desde 2009: uma em 2009; sete em 2010; três em 2011; uma em 2012; três em 2013; uma em 2015. Podemos observar que a difusão da PII se deu de forma lenta e bastante limitada em quantidade. Além disso, nenhuma universidade apresentou uma PPD, o que nos remete aos achados de autores como Ribeiro (2012) e Medeiros e Ferreira (2014), os quais concluem que a cultura de preservação digital nas instituições federais de ensino superior ainda é escassa.

Os RIs são a via verde, referida pela *Budapest Open Access Initiative*, para o acesso aberto à literatura científica (BOAI, 2002). As universidades, quando criam um RI, não apenas permitem o acesso às suas produções intelectuais, mas também tencionam preservar, a longo prazo, o conteúdo do acervo. Observamos que várias universidades que não estabeleceram uma PII declaram essa intenção na *homepage* dos seus RIs. Conforme nos ensina Suber (2015), tais repositórios podem armazenar artigos de periódicos, teses e dissertações, materiais usados em cursos, arquivos de áudio e vídeo, material digitalizado etc. Por conseguinte, as partes interessadas (os chamados *stakeholders*) dos RIs demandam um programa para guiar suas atividades, no sentido de preservar esses materiais a longo prazo. Acrescente-se que esse programa deveria ser considerado como um compromisso institucional declarado por meio de uma política de preservação digital.

Algumas universidades, apesar de terem feito o esforço de aumentar a visibilidade de seus RIs cadastrando-os no OpenDOAR, não parecem dar importância à ferramenta de políticas provida por esse diretório. Da mesma forma, pelo que se depreende das PIIs, as pretensões de preservação nem sequer se desdobraram na elaboração de uma política para este fim. Assim, para mover-se do interesse de preservar a execução de um programa, uma política de preservação digital deveria ser planejada pelo *staff* dos RIs em parceria com os administradores das comunidades e subcomunidades. Uma tal política, necessariamente, deve ser aprovada pelo Conselho Universitário a fim de oficializar o compromisso institucional com as questões relativas à preservação digital. Uma política bem estruturada depende do investimento em capacitação do *staff* do repositório. Neste sentido, as universidades poderiam estabelecer parcerias educacionais com o IBICT para treinar o pessoal envolvido no planejamento e nas práticas da preservação digital.

A análise documental evidencia a consciência das universidades federais sobre a importância da preservação digital, mas não é suficiente para revelar a existência (ou a

não existência) de práticas de preservação nessas instituições e os demais elementos que podem dar suporte à elaboração de uma política de preservação digital. Sendo assim, um outro levantamento se fez necessário para estudarmos a cultura de preservação digital em todos os RIs pesquisados.

5.2 – Questionário feito aos administradores dos repositórios institucionais

Em uma PPD, a declaração de princípios é a enunciação de um conjunto de compromissos de uma instituição com a preservação do acervo do seu repositório. Conforme o InterPARES (2012), um desses compromissos estabelece os princípios gerais que orientam o gerenciamento e a prática da preservação dos objetos digitais assegurando a confiabilidade, autenticidade e acessibilidade desses objetos ao longo do tempo. Neste sentido, a política do UK Data Archive presta-se como um exemplo, quando, entre seus propósitos, afirma:

> (...) To ensure the continued use of these resources the Archive follows a policy of active preservation with the aim of ensuring the *authenticity*, *reliability* and *logical integrity* of all resources entrusted to its care while providing formats suitable for research, teaching or learning, in perpetuity (UK Data Archive, 2014, p. 3).

Dentre as 63 universidades federais, 25 (40%) nem sequer possuem um repositório. Dos 38 repositórios existentes, 16 apresentam uma política de informação institucional. Mesmo assim, algumas PIIs publicadas não estão assinadas pelo reitor ou por outra autoridade administrativa, o que compromete a validade jurídica desses documentos. Além disso, nenhuma universidade construiu uma PPD para seus repositórios.

Nessas circunstâncias, havemos de questionar onde poderá florescer um compromisso com a preservação digital nos repositórios das universidades em análise. Possivelmente, o ponto de partida está na consciência do *staff* dos repositórios, porque essa é a parcela dos *stakeholders*, a qual pode liderar a elaboração de uma PPD e implementá-la. Por essa razão, a nossa primeira questão procurou identificar o grau de conscientização sobre a falta de uma política de preservação entre os administradores dos RIs. Conforme podemos visualizar no gráfico 1 a seguir, 47% dos respondentes (superadministradores dos repositórios) afirmaram que os administradores das

comunidades, subcomunidades e coleções reclamam a falta de uma PPD contra 42%, que afirmaram não haver esse tipo de reclamação. Em suas *homepages*, vários RIs declararam o objetivo de preservar o material depositado, sem contudo fazer referência a qualquer procedimento de preservação. Isso pode indicar que os administradores entendem que as funcionalidades de preservação do *Dspace* serão suficientes para garantir o acesso a longo prazo. Apenas 11% dos respondentes não souberam informar se os administradores das comunidades e coleções reclamam, ou não, a falta de uma PPD. O elevado percentual de respondentes, que afirmaram não haver queixa quanto à falta de uma política de preservação, sugere a necessidade de os *stakeholders* compreenderem a complexidade subjacente à preservação digital e serem conscientizados sobre a importância de uma PPD, para o êxito do acesso aos objetos digitais a longo prazo.

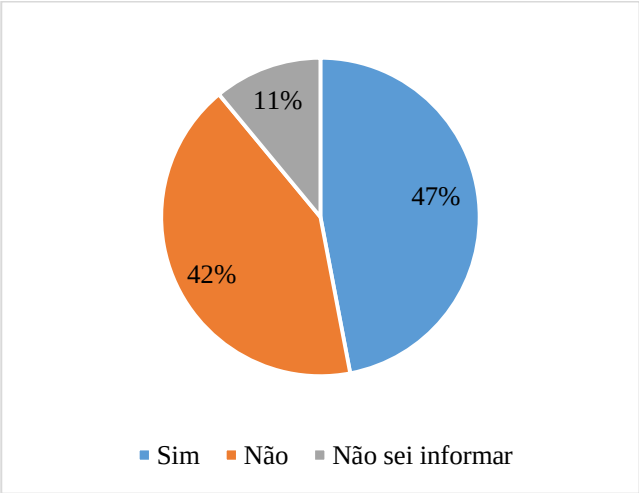

Gráfico 1 – Reclamação quanto a falta de uma PPD (N=38)

Fonte: dados da pesquisa, 2017.

No *framework* do InterPARES (2012), o elemento de política "*related sources*" recomenda referenciar outras políticas da instituição, ou até políticas externas, que possam ser relacionadas com uma PPD; porém não explica o porquê deste escrutínio nem o modo com que as intersecções entre as políticas ocorrem. Sheldon (2013) elaborou uma taxonomia para destacar os elementos comuns a 33 PPDs encontradas em instituições de várias partes do mundo, mas não detectou um elemento que relacionasse a preservação digital com as demais estratégias das instituições pesquisadas. *The Ohio State University Libraries* estruturou seu *framework* com base nos elementos comuns a uma dúzia de

outras PPDs (Noonan, 2014). O primeiro deles, "*introduction or purpose*", deverá contextualizar e articular a necessidade da política.

Beagrie *et al*. (2008) ensinam que o levantamento das políticas e estratégias organizacionais existentes é a primeira tarefa a ser executada quando se está planejando a elaboração de uma PPD. Procura-se com isso mapear a relação entre as estratégias da organização e a preservação digital, porque uma política para este fim só fará sentido se for alinhada com o núcleo das políticas que dirigem os negócios e estratégias de uma instituição. Por exemplo: conforme explicam esses autores no apêndice 4 (*Information Strategies*), seção *Information Security*, uma universidade deveria precaver-se contra ameaças de segurança aos seus ativos informacionais de modo que evitasse problemas operacionais e relativos à sua própria reputação. Esses problemas podem estar relacionados com a disponibilidade (garantia do acesso permanente à informação), com a integridade (certeza de que a informação é confiável e atualizada) ou com a confidencialidade (proteção das informações sensíveis contra o uso não autorizado). Uma política de segurança universitária deveria ser estruturada de modo que atribuísse os papéis e responsabilidades para com a segurança da informação. Por conseguinte, a ligação de uma política de segurança com uma PPD garante que a informação sempre esteja acessível.

As universidades federais brasileiras também possuem estratégias para orientar sua atuação no âmbito da administração e da informação. Na área administrativa, o Plano de Desenvolvimento Institucional (PDI) define a visão, missão e valores institucionais, os objetivos estratégicos, diversas políticas institucionais etc. Um exemplo é o PDI 2014-2018 da Universidade Federal da Paraíba (UFPB, 2014). Esta universidade também promulgou uma política de segurança da informação (UFPB, 2014b). Como todas as demais universidades federais, disponibiliza um *guideline* (UFPB, 2016) para o depósito de teses e dissertações, acompanhado de um termo de autorização fundamentado na Lei nº 9610/98, que regula os direitos autorais. Conforme foi dito, 16 universidades promulgaram uma política de informação institucional. Diante desse panorama, procuramos saber quais as políticas e estratégias institucionais poderiam ser relacionadas com uma PPD. Como ficou demonstrado no quadro 8, adiante, 63% dos respondentes assinalaram a PII do próprio repositório. Neste caso, observa-se, por exemplo, que o artigo 9º da PII da Universidade Federal de Ouro Preto (UFOP, 2013) prevê o

estabelecimento de outras políticas ou procedimentos específicos, a fim de garantir o pleno povoamento do repositório. Na sequência desse levantamento, 39% assinalaram a política de informação da universidade; 39%, o plano de desenvolvimento institucional; 34%, a política de segurança da informação da universidade e 8% não souberam informar. No campo outros do questionário, um respondente indicou o plano diretor de tecnologia da informação e um segundo respondente aparentemente não entendeu a pergunta e respondeu que não havia uma PPD. Esse campo concentrou 5% das respostas.

Quadro 8 – Políticas institucionais associadas com uma política de preservação digital

Tipos de Políticas Institucionais Utilizadas pelos Ris	**Percentual de Uso**
A política institucional de informação do próprio repositório	63%
A política de informação da universidade	39%
A Política de Desenvolvimento Institucional (PDI)	39%
A política de segurança da informação da universidade	34%
Não sei informar	8%
Outras	5%

Fonte: dados da pesquisa, 2017.

Como vimos na seção 2.8 (Acesso), o modelo OAIS define a comunidade-alvo (*Designated Community*) como um grupo de usuários potenciais aptos a entender um conjunto de informações específicas. A comunidade-alvo pode ser constituída de múltiplas comunidades de usuários. Ela é definida pelo repositório. Essa definição pode mudar ao longo do tempo (CCSDS, 2012). Conhecer a comunidade-alvo influi na qualidade do serviço de um repositório e torna usável e inteligível a informação preservada. Esse é o objetivo da preservação. A comunidade-alvo pode diferir por coleções e assim ser definida pelo tipo de coleção (Sierman *et al*, 2014).

Todos os repositórios das universidades federais utilizam o *DSpace*. Esta plataforma foi projetada em conformidade com o modelo de referência OAIS (Tansley, Bass e Smith, 2003), ou seja, os repositórios estão organizados em comunidades e coleções. O RI da UFRGS, por exemplo, define a estrutura organizacional do repositório na sua PII (UFRGS, 2010, p. 1):

> Art. 3º − O Repositório está organizado em Comunidades, Subcomunidades e Coleções. As comunidades e suas subdivisões são grupos que fornecem

> conteúdos para o portal. As coleções são conjuntos de itens, aos quais estão associados metadados e objetos digitais. Os objetos digitais podem conter texto, imagem, vídeo e áudio.

Os objetos digitais depositados nos repositórios são de acesso livre, mas o uso desses objetos deve ser regulado por licenças específicas e sempre com base na legislação do direito autoral brasileiro. Assim, o RI exemplificado anteriormente dedica três artigos de sua PII para definir os critérios de acesso e uso. São eles:

> Art. 13º − Os trabalhos depositados no Lume estão disponíveis gratuitamente para fins de pesquisa e estudo de acordo com a licença pública *Creative Commons* adotada no Lume;
>
> Art. 14º − O autor é titular dos direitos autorais dos documentos disponíveis no repositório, é vedado, nos termos da lei, a comercialização de qualquer espécie sem sua autorização prévia;
>
> Art. 15º − Os usuários que utilizarem qualquer trabalho, no todo ou em partes, em novas publicações ficam obrigados a citá-lo, indicando o nome do autor e os dados completos da obra (UFRGS, 2010, p. 3-4).

Conforme nos ensinam Sierman *et al.* (2014), a reputação de uma instituição está associada a uma política que enumera os direitos importantes e descreve como a instituição irá tratá-los. Não obstante, os direitos de acesso podem ser aplicados apenas a uma parte da comunidade de usuários (por exemplo: usuários de uma determinada coleção) ou a uma coleção específica. Neste sentido, coube-nos indagar se haveria algum tipo de documento cujo acesso é restrito aos membros da universidade. Resultado, 53% deram resposta afirmativa e 47% resposta negativa (gráfico 2).

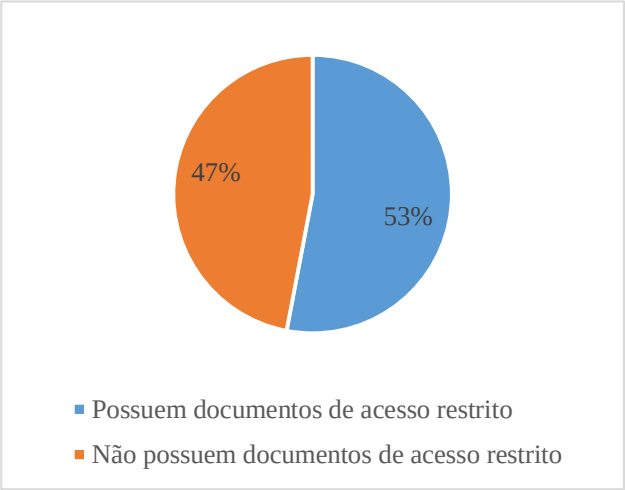

Gráfico 2 – RIs que possuem, ou não, documentos de acesso restrito (N=38)

Fonte: dados da pesquisa, 2017.

O tipo de conteúdo aceito por um repositório deverá ser discriminado a fim de orientar os depositantes e os administradores das coleções. Por exemplo: o RI Dspace@MIT declara aceitar todo tipo de conteúdo digital e lista alguns exemplos na seção *FAQ*, tais como: artigos, *preprints*, teses, arquivos de áudio e vídeo, publicações multimídias etc. Complementando essas informações, a seção *Policies – Content Guidelines* − lista as seguintes características dos trabalhos a serem depositados:

> - O trabalho deve ser produzido, submetido e patrocinado pelas faculdades do MIT.
> - O trabalho deve ser orientado à educação ou à pesquisa.
> - O trabalho deve estar em formato digital.
> - O trabalho deve estar completo e pronto para distribuição.
> - O autor ou proprietário deve estar disposto e apto para conceder ao MIT o direito de preservar e distribuir o trabalho via DSpace@MIT.
> - Se o trabalho for parte de uma série, outros trabalhos na série também deveriam ser adicionados de modo que o DSpace@MTI possa oferecer um conjunto tão completo quanto seja possível (DSpace@MIT, 2016, tradução nossa).

O RI da Universidade do Minho (RepositóriUM, 2014) é um outro exemplo importante devido ao papel relevante no movimento de acesso aberto desempenhado por essa universidade. O RepositóriUM responde na seção *FAQ* quais são as características obrigatórias dos documentos a serem depositados:

> No RepositóriUM podem ser depositados qualquer tipo de documento, em qualquer formato, desde que reúna as seguintes condições básicas:
>
> - Ser produzido (autor ou co-autor) por membro (s) da UMinho
> - Não ser efémero
> - Estar em formato digital
> - Estar completo (texto integral) e pronto para "publicação".
> - O autor deve poder, e estar disposto a, conceder à UMinho o direito não-exclusivo de preservar e dar acesso ao seu trabalho através do RepositóriUM.

Adicionalmente, a Universidade do Minho estabelece uma tipologia documental específica mediante sua política de autoarquivo. Esta determina aos docentes e pesquisadores depositar uma cópia de seus artigos científicos, dos trabalhos apresentados em congresso, das conferências e de outros textos científicos. Essa política inclui, ainda, as teses e dissertações produzidas pelo seu corpo discente.

Os RIs das universidades federais não informam, em suas respectivas *homepages*, quais são os tipos de documentos aceitos nem tampouco disponibilizam uma lista do tipo *FAQ*. A exceção fica por conta do RI da UFBA e do RI da UFT. O primeiro disponibiliza

em sua *homepage* um fôlder intitulado "Orientações para uso do Repositório Institucional da UFBA" (UFBA, 2010). Nesse fôlder, são listados os tipos de documento aceito, em quatro categorias: produção bibliográfica, produção técnica, produção cultural e trabalhos finais e parciais de cursos. O segundo também disponibiliza um fôlder para apresentação do repositório intitulado "Repositório Institucional e Biblioteca Digital de Teses e Dissertações da UFT" (UFT, 2011), no qual são listados todos os tipos de conteúdo que constituem o acervo do repositório.

Entre os repositórios que utilizam, integralmente ou parcialmente, o modelo de PII do IBICT, a tipologia documental é restrita aos artigos publicados em periódicos científicos e aos trabalhos que foram submetidos a uma banca de especialistas. Apenas as PIIs da UFT e da UFRGS definem os requisitos dos documentos de modo semelhante ao Dspace@MIT e ao RepositóriUM. Neste sentido, merece destaque o seguinte artigo da PII da UFRGS:

> Art. 9º Para ser incluído em Comunidades e Coleções do LUME, o objeto digital deve atender aos requisitos gerais relacionados a seguir:
>
> I – ser produzido ou orientado por membro(s) da UFRGS;
>
> II – não ser efêmero;
>
> III – ser de acesso livre;
>
> IV – estar em formato digital, conforme definido pela equipe técnica do Lume;
>
> V – estar completo e finalizado;
>
> VI – conter metadados e objeto(s) digital(is);
>
> VII – o autor deve poder estar disposto a conceder à UFRGS o direito não exclusivo de dar acesso ao público pela Internet e de preservar seu trabalho integral no Lume;
>
> VIII – o autor deve ter obtido o direito de reprodução de conteúdos criados por outros, mas que façam parte de seu trabalho (UFRGS, 2010).

Esse artigo deixa implícito no inciso I que o repositório aceita qualquer tipo de documento e, consideradas as outras condições nos incisos subsequentes, poderíamos supor que ao depositante não restariam dúvidas. Entretanto, como vimos pela disposição do Dspace@MIT e do RepositóriUM em ajudar os usuários, a existência de uma seção para *FAQs* é fundamental para orientar os usuários do sistema. Como os RIs das universidades federais brasileiras não oferecem esse tipo de informação, perguntamos quais são os tipos de documento que os repositórios aceitam para preservar e obtivemos a seguinte lista:

- Artigos de periódicos científicos.
- Dissertações.
- Teses.
- Livros.
- Capítulos de livros.
- Trabalhos de conclusão de curso.
- Relatórios técnicos.
- Anais de eventos científicos.
- Vídeos de conferências.

Nem todos os repositórios incluíram relatórios, anais e imagens, mas alguns deles responderam que aceitam toda a produção intelectual das suas respectivas comunidades acadêmicas. Todavia, em um acervo grande e heterogêneo, há que se fazer uma seleção do que pode ser preservado, sob pena de se aumentar o custo financeiro e operacional de um programa de preservação digital. Saliente-se, contudo, a importância de se caracterizar o conteúdo do acervo de um RI para se dar clareza sobre a abrangência de uma PPD, conforme podemos depreender das leituras em Beagrie *et al*. (2008) e no módulo 2 do InterPARES (2012)

Uma instituição deve declarar ser responsável pela legibilidade e integridade técnica dos objetos digitais a longo prazo. Como vimos na seção 2.4, o objeto digital pode ser recebido em seu formato original ou copiado. Contudo, poderá sofrer diversas ações de preservação durante o seu ciclo de vida, a fim de se manter acessível e autêntico. Uma dessas ações pode ser a migração de formatos de arquivo. Visto que os repositórios podem aceitar qualquer formato de arquivo, uma política de preservação deve incluir um acordo com os autores para migrar uma cópia do objeto depositado para um formato destinado à preservação. Neste sentido, na questão 5 do questionário aplicado para esta pesquisa, perguntamos se os repositórios estabeleciam um acordo com os depositantes para realizar uma conversão de formato de arquivo para fins de preservação e acesso. O gráfico 3 mostra que 63% dos repositórios não realizam esse tipo de acordo. Apesar de não existir uma PPD, 37% dos respondentes afirmaram haver tal tipo de acordo.

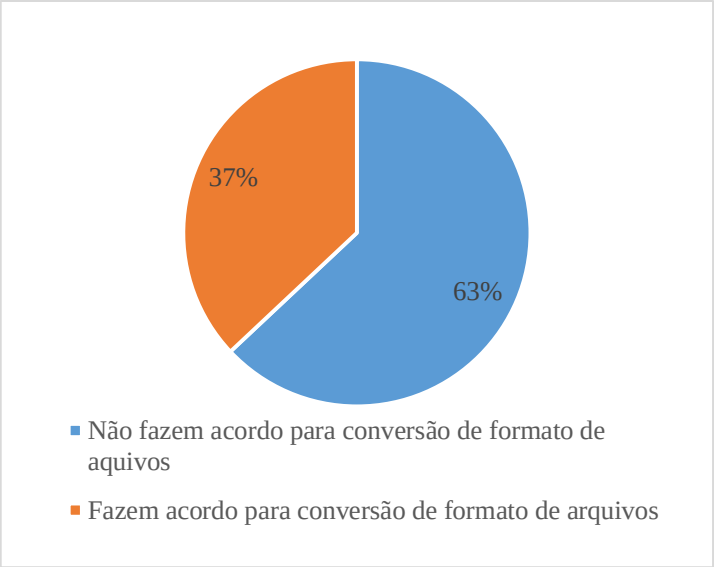

Gráfico 3 – RIs que fazem, ou não, acordo de conversão de formato de arquivo (N=38).

Fonte: dados da pesquisa, 2017.

Ainda na seção 2.4, vimos que um objeto − ou uma coleção − poderá ser descartado a qualquer tempo por diversas razões, inclusive as de ordem legal. Por conseguinte, uma instituição deverá adotar um gerenciamento de descarte, além de informar qual será o destino do material removido. Neste sentido, indagamos se os repositórios definem critérios de exclusão de documentos (objetos) ou de coleções. Como ficou demonstrado no gráfico 4, adiante, 63% responderam que não definem critérios de exclusão para documentos ou coleções, 24% responderam que definem para ambos e 13% definem apenas para os documentos. A opção de resposta ‘sim, apenas para coleções’ não foi assinalada. Estes resultados indicam que a maioria dos repositórios está correndo o risco de perder sua respectiva credibilidade, caso algum material tenha que ser removido.

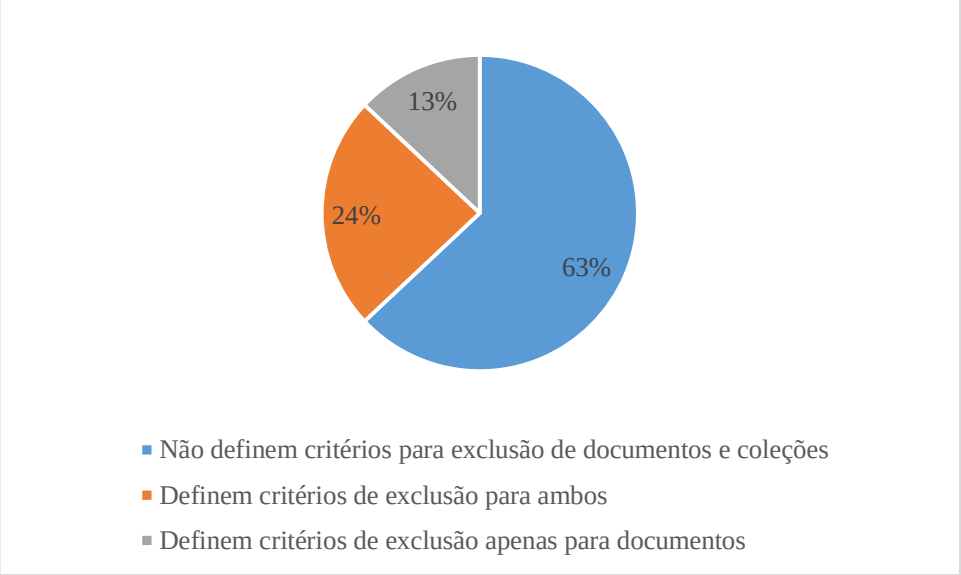

Gráfico 4 – Percentual dos RIs que fazem, ou não, acordo de exclusão do material digital (N=38).

Fonte: dados da pesquisa, 2017.

Nas políticas de preservação digital de RIs do mundo todo, é comum haver um glossário a fim de facilitar a compreensão do jargão científico e de termos específicos. A criação de um glossário também é uma recomendação de todos os *frameworks* para a criação de uma PPD. Tendo em vista que exploramos as práticas de preservação digital das universidades pesquisadas, procuramos saber se existiria um glossário de preservação digital. Como se pode ver no gráfico 5, 95% responderam que não possuem um glossário, contra 5% que afirmaram possuí-lo. Até este ponto da pesquisa, tínhamos a evidência de que não existia PPD para os RIs das universidades federais, porque simplesmente não há qualquer documento com esse conteúdo nos repositórios pesquisados. Mesmo assim, esse resultado reforça o indício da inexistência desse tipo de política, isto é, ele corrobora a nossa hipótese de que as universidades não produziram esse tipo de política.

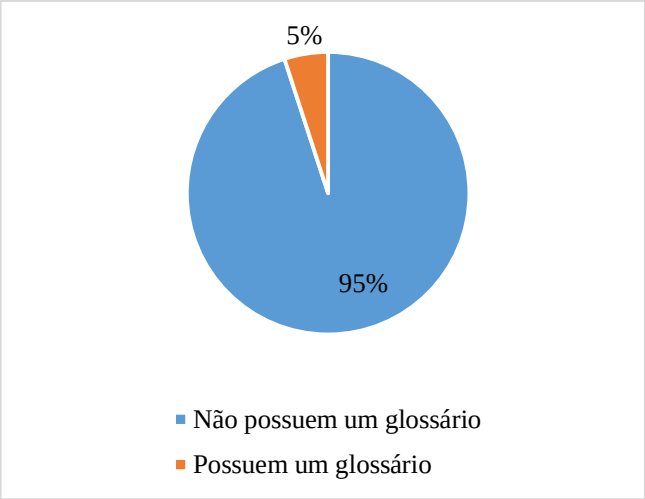

Gráfico 5 – RIs que possuem, ou não, um glossário de preservação digital (N=38)

Fonte: dados da pesquisa, 2017.

Conforme a proposta de Beagrie *et al*. (2008), uma política de preservação digital deve apresentar os responsáveis pela operacionalização e financiamento da preservação dentro da instituição. A ferramenta de política da ERPANET (2003) acrescenta que essa responsabilidade deve ser atribuída a setores da estrutura organizacional. Isto significa nomear os setores e equipes que possuem a incumbência de viabilizar um programa de preservação digital. Por exemplo: a política de preservação da biblioteca da Yale University Library (YUL, 2015) estabelece que o Departamento de Preservação é o responsável pela administração dessa política e pela alocação dos recursos financeiros para a preservação de todos os materiais analógicos e digitais.

Dentre as 38 universidades federais, 53% afirmaram possuir um setor responsável pela preservação digital e 47% responderam que não possuem um tal setor (gráfico 6).

Gráfico 6 – Universidades possuidoras ou desprovidas de um setor de preservação digital (N=38)

Fonte: dados da pesquisa, 2017.

Quando se perguntou quais eram os setores ou equipes responsáveis pela preservação do acervo dos RIs, 58% dos respondentes nomeou algum setor, isto é, 24% indicaram a Biblioteca Central, que tem parceria com a Superintendência (ou Núcleo) de Tecnologia da Informação, a Superintendência (ou Núcleo) de Tecnologia da Informação foi indicada por 18% dos respondentes e 16% indicaram a Biblioteca Central. No campo outros do questionário, concentraram-se 21% das respostas. Dentre elas, quatro indicaram não haver tais setores e quatro nomearam, respectivamente, os setores Centro de Computação e *Software* Livre (grupo de pesquisa), Coordenação de Bibliotecas (esta tem a mesma competência administrativa de uma biblioteca central), Diretoria de Governança Informacional e Diretoria de Tecnologia da Informação (esses dois últimos equivalem a uma Superintendência (ou Núcleo) de Tecnologia da Informação. Por fim, 21% afirmaram não haver setores ou equipes com essa incumbência (gráfico 7). Esse resultado demonstra que poucas universidades realizam um trabalho multidisciplinar nesse campo, visto que apenas 24% dos repositórios compartilham a responsabilidade com a preservação entre as bibliotecas centrais e os núcleos de tecnologia da informação. As demais instituições precisam ser conscientizadas sobre a imprescindível necessidade de constituírem-se equipes multidisciplinares, pois a complexidade para planejar e operacionalizar um programa de preservação digital exige um *staff* com formações e áreas de atuação diferenciadas.

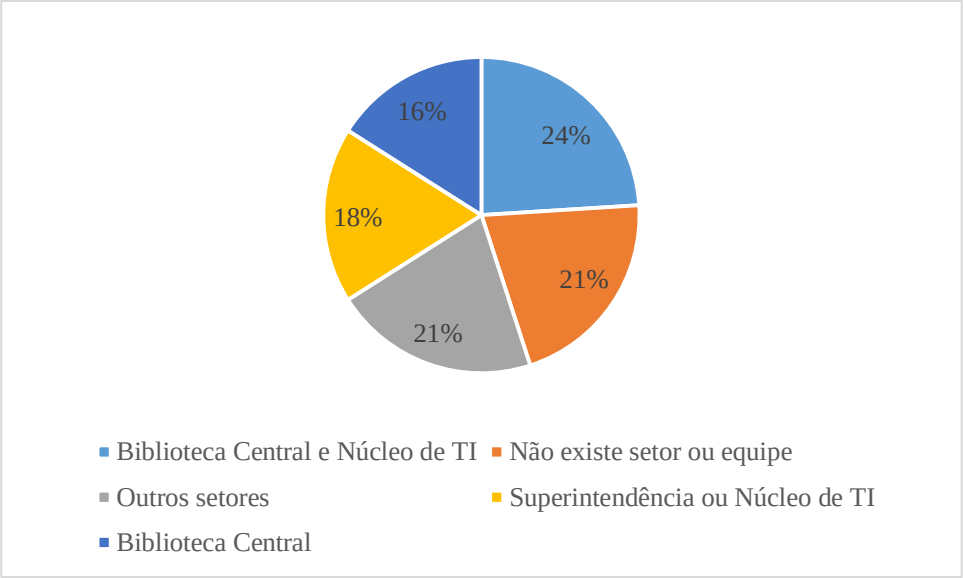

Gráfico 7 – Setores responsáveis pela preservação digital (N=38)

Fonte: dados da pesquisa, 2017.

A sustentabilidade financeira é um dos sete atributos que caracterizam um repositório digital confiável (Beagrie *et al.*, 2002). Por conseguinte, uma organização universitária deve declarar de onde sairão os recursos destinados a fomentar a preservação digital em seus RIs. Conforme explica o *framework* elaborado por McGovern (2007), que é um modelo adotado pelas políticas de preservação digital do *Inter-university Consortium for Political and Social Research* e do *Canadian Heritage Information Network*[54], a instituição, ao declarar um compromisso financeiro com um programa de preservação digital confirma e resume o apoio dado ao programa e informa de onde virão os recursos para sustentar as ações de curadoria e preservação digital.

Apesar de as universidades pesquisadas nomearem os setores responsáveis pela atividade de preservação digital, os RIs não fazem menção a qualquer fonte de financiamento, em suas *homepages*, ou na seção de documentos sobre o próprio repositório. Considerando que essa atividade envolve o investimento em infraestrutura tecnológica e a formação de pessoal, perguntamos aos administradores qual setor (ou setores) poderia planejar a destinação de recursos para um programa de preservação digital. A maioria dos respondentes indicou mais de um setor e dois não souberam dizer se existe algum. Assim, observamos a incidência de 20 indicações para as bibliotecas

[54] Disponível na WWW: http://canada.pch.gc.ca/eng/1454520330387

centrais, 18 para os setores de gestão de TI, 4 para as pró-reitorias de planejamento, 3 para as pró-reitorias administrativas, 6 para outros setores, quais sejam: reitoria, vice-reitoria, pró-reitoria de pós-graduação, pró-reitoria de pesquisa, criação e inovação, arquivo e editora universitária (quadro 9). Contudo, constatamos um caso excepcional: uma universidade referiu a existência de um planejamento financeiro para suas atividades de preservação digital, elaborado pela Biblioteca Central e Pró-Reitoria de Comunicação, Informação e Tecnologia. Ambos fazem parte da estrutura da administração superior, sendo a biblioteca central um órgão suplementar e a pró-reitoria um órgão auxiliar. Aos órgãos suplementares compete o apoio didático, científico e tecnológico a determinados setores ou a toda a universidade. Normalmente, a um órgão suplementar é destinada uma pequena parcela do orçamento de uma universidade federal. Nem sempre essa parcela é suficiente para custear nem sequer as despesas de manutenção desse tipo de setor.

Quadro 9 – Setores indicados para dar sustentabilidade financeira aos RIs

Setor	**Número de Indicações**
Biblioteca Central	20
Órgãos de gerenciamento de TI	18
Pró-Reitoria de Planejamento	4
Pró-Reitoria Administrativa	3
Outros	6

Fonte: dados da pesquisa, 2017.

No modelo de política proposto por Beagrie *et al.* (2008), a propriedade intelectual é uma cláusula que demonstra como a instituição planeja reconhecer e tratar as questões de *copyright*. Dessa forma, a instituição deverá estabelecer o contexto legal que permita a reprodução de um objeto digital com a finalidade de preservá-lo, bem como acordos e métodos de depósito, isto é, autodepósito, depósito mediado pelo *staff* etc.

Na verdade, as políticas de direitos de propriedade intelectual são um desafio para um programa de preservação digital, porque elas devem cumprir a lei, a despeito da possibilidade de impactar negativamente o processo de preservação digital (CHIN, 2016). Por essa razão, elaboraram-se propostas para minimizar os conflitos de interesse entre os direitos de propriedade intelectual e as práticas de preservação digital, conforme

abordamos no final da seção 2.6 – Propriedade Intelectual. O *Digital Preservation Handbook* relaciona diversos procedimentos para negociações de direitos entre os depositantes, detentores de direitos e os repositórios comprometidos com a preservação digital. Dentre eles, destacamos o que recomenda o desenvolvimento de uma carta-modelo para a compensação de direitos, modelos de acordos de depósito, modelos de licença e cláusulas para as atividades de preservação (DPH, 2015).

O quadro 10, a seguir, exibe os tipos de licença relacionada com os direitos autorais aplicados no ato da submissão de documentos nos RIs das universidades federais. A licença *Creative Commons* é adotada por 79% dos repositórios, a declaração de distribuição não exclusiva é adotada por 39% deles, 8% afirmaram não aplicar qualquer tipo de licença e 10% afirmaram utilizar a licença padrão da universidade, isto é, um termo de autorização em que o autor permite que seu trabalho seja publicado no repositório. Esse termo deve incluir uma declaração de licença não exclusiva e uma licença de uso com base na licença *Creative Commons* e na Lei 9.610/98, que regula o direito autoral. Porém, dos treze repositórios que disponibilizam o termo de autorização em suas *homepages*, apenas 6 foram elaborados com todos os elementos. Pode-se encontrar um termo de autorização que não faz referência a nenhuma licença ou mesmo à Lei 9.610/98.

Quadro 10 – Licenças utilizadas pelos RIs das universidades federais

Tipo de licença	**Percentual de uso pelos Ris**
Licença *Creative Commons*	79%
Declaração de distribuição não exclusiva	39%
Termo da autorização	10%
Nenhuma licença é aplicada	8%

Fonte: dados da pesquisa, 2017.

De todo modo, nenhum termo de autorização encontrado nos repositórios contém uma licença específica para a preservação digital. Todavia, como se pode visualizar no gráfico 8, 26% dos respondentes afirmaram que as licenças dos seus repositórios incluem uma cláusula relativa à preservação do material que irá ser depositado e 74% afirmaram não existir tal cláusula.

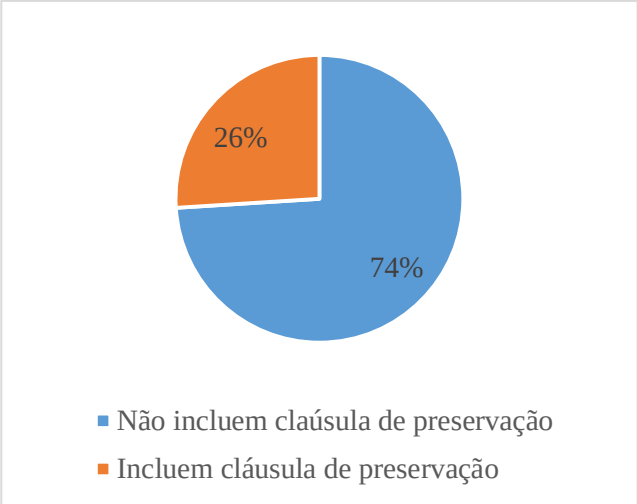

Gráfico 8 – RIs incluidores, ou não, de uma cláusula de preservação digital (N=38)

Fonte: dados da pesquisa, 2017.

A preservação de *bit* requer, dentre outros procedimentos, a decisão sobre o número de cópias, a distribuição geográfica e a distribuição organizacional. Dessa forma, minimiza-se o risco da perda de dados criando-se cópias do material digital em mídias diferentes, armazenadas em diferenciadas localizações geográficas e administradas por equipes ou instituições diferentes (Sierman, Jones e ElstrØm, 2014). Um exemplo desse tipo de procedimento foi estabelecido no documento de estratégias de preservação da instituição dinamarquesa *State and University Library* (SUL, 2016, p. 9):

> As a minimum, the State and University Library keeps at least two copies of data. One of these is kept at the library's main address and the other copy is located at the library's facilities in Skejby, Aarhus. The two copies will be stored by using different technologies, and the library makes sure that both copies are not controlled by the same organisational unit and/or person.

Os níveis de preservação propostos pela NDSA (seção 2.2 – Preservação de *bit*), que tratam da redundância geográfica, também são indicados pelo *Digital Preservation Handbook* (DPH, 2015), com a recomendação de se fazer uma combinação de cópias *online* e *off-line*. Esse *handbook* relaciona os mais notáveis sistemas de armazenamento para preservação digital. Dentre eles destacamos o *DSpace* e o *Lockss*. Todos os 38 RIs das universidades federais utilizam o *DSpace* como plataforma de armazenamento e preservação. O *Lockss* é utilizado por 7 universidades (UFSM, UFRGS, UFG, UNB, UFBA, UFPB, UFRN), que são parceiras integrais da Rede Cariniana (Rede de Serviços de Preservação Digital) do IBICT (2015). Essas universidades já armazenam todos os seus periódicos que estão na plataforma *Open Journal System* na Rede Cariniana. Esta

pretende incluir também as teses e dissertações que estão nos repositórios. Assim, procuramos saber a posição das universidades em face da inclusão dos repositórios na Rede Cariniana. Conforme demonstra o gráfico 9, 55% pretendem incluir o repositório na Cariniana, 29% dos respondentes declaram as seguintes respostas no campo outros:

- "Isto não foi decidido."
- "Não existe nenhum planejamento no que diz respeito a qualquer Serviço de Preservação Digital."
- "Ainda não pensamos a respeito."
- "Estamos nos primeiros contatos iniciais."
- "Em estudo."
- "Atualmente estamos tentando realizar uma migração do RI para uma biblioteca digital que incluirá outros trabalhos. A partir daí será solicitada inserção na Rede."
- "Como o repositório atravessa vários problemas técnicos devido ao próprio setor de TI, nesse momento o repositório não está inserido no âmbito de nenhuma rede, tampouco seria interessante tal inclusão nesse momento."
- "Sim, no momento não estamos incluídos em nenhuma rede de preservação Digital, porém , pretendemos fazer parte da Rede Cariniana."
- "Pretende solicitar inclusão na rede Cariniana."
- "Não tinha conhecimento da Rede."
- "A [Universidade] é uma caixa *Lockss* e teremos o repositório na Rede."

Os demais 11% afirmaram que o repositório está parcialmente incluído na Cariniana e 5% garantiram que o repositório está totalmente incluído na Rede Cariniana.

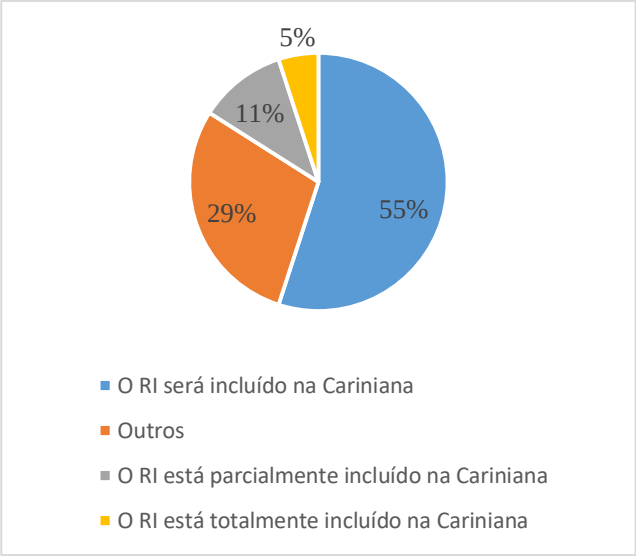

Gráfico 9 – Inclusão dos RIs na Rede Cariniana (N=38)

Fonte: dados da pesquisa, 2017.

A compatibilidade com o modelo de referência OAIS é o primeiro atributo de um repositório confiável. Em geral, asseguram-na, as políticas de preservação digital, sendo esse modelo um padrão para os repositórios de acesso aberto. Um exemplo pode ser lido na política de preservação digital do ICPSR (2012, p. 1):

> In achieving its digital preservation objectives, ICPSR recognizes the need to comply with the prevailing standards and practice of the digital preservation community. ICPSR is committed to developing its digital preservation policies, repository, and strategies in accordance with the Open Archival Information System (OAIS) Reference Model (2012). ICPSR tracks and responds to related OAIS initiatives, including developments in digital archives certification, persistent identifiers, preservation metadata, and the producer-archive interface. The mapping of ICPSR's preservation process to OAIS is synthesized in Digital Preservation Requirements Applied to ICPSR.

A plataforma *DSpace* é compatível com o modelo OAIS e mantém três tipos de metadado: descritivo, administrativo e estrutural. Dessa forma, os padrões Dublin Core, METS e PREMIS são suportados por esse sistema. Quando uma instituição pratica alguma atividade de preservação digital além da utilização de metadados descritivos, espera-se que alguns membros do *staff* do repositório estudem a visão geral das funcionalidades do *DSpace* e, consequentemente, cheguem ao conhecimento dos demais esquemas de metadados. No entanto, nenhum respondente de nossa pesquisa assinalou uma das opções para indicar se utilizam o METS ou o PREMIS; 84% utilizam o *Dublin Core*, exclusivamente; 13% utilizam o *Dublin Core* e o OAIS e 3%, apenas o OAIS (gráfico 10).

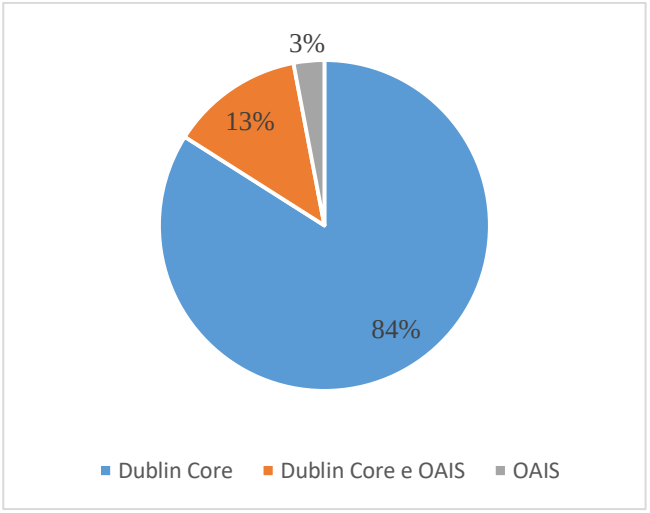

Gráfico 10 – Padrões utilizados pelos RIs (N=38)

Fonte: dados da pesquisa, 2017.

Na seção 2.7 – Padrões –, descrevemos os principais membros das famílias de padrões mais adotados para a preservação digital, isto é:

- Os métodos de auditoria para repositórios digitais.
- A descrição de formato de arquivos.
- Os metadados de preservação

O modelo de referência OAIS é básico para os demais padrões que surgiram na esteira da preservação digital e dos repositórios digitais. O METS e o PREMIS são membros da família de padrões de metadados de preservação.

Os formatos de arquivos estão sujeitos à obsolescência tecnológica e constituem um desafio para a preservação digital. Seja um arquivo proprietário com especificações fechadas ou abertas, seja um arquivo não proprietário com especificações abertas, todos possuem seu grau de vulnerabilidade causada pelos interesses e demandas do mercado de *software*. Por essa razão, o tutorial *Digital Preservation Management: Implementing Short-Term Strategies for Long Term Problems*, na seção 3.1 – *Obsolescence: File Formats and Software* (Kenney *et al.*, 2014), recomeda as seguintes ações de preservação, aqui relacionadas por tópicos: fazer a normalização dos arquivos; utilizar o PRONOM (pertencente à família de padrões de descritores de formato de arquivos) e a estratégia de emulação, quando isto for o caso, além da consulta a diversos artigos científicos referenciados pelo tutorial.

Todavia, é no tópico *Format Support* da seção *Policies* no repositório DSpace@MIT (2016) que encontramos um exemplo de uma política para formatos de arquivos. Ela estabelece três categorias de suporte para os formatos: *supported, known, unsupported*. Essas categorias são utilizadas na coluna *level* da tabela que define a coleção de formatos de referência (*Format Reference Collection*). Os 38 repositórios da universidades federais brasileiras não apresentam nem sequer uma *guideline* que se aproxime desse tipo de política. Por outro lado, 53% dos respondentes afirmaram que os seus respectivos repositórios aceitam formatos abertos e proprietários, 45% aceitam apenas formatos abertos e 2% aceitam apenas formatos proprietários (gráfico 11).

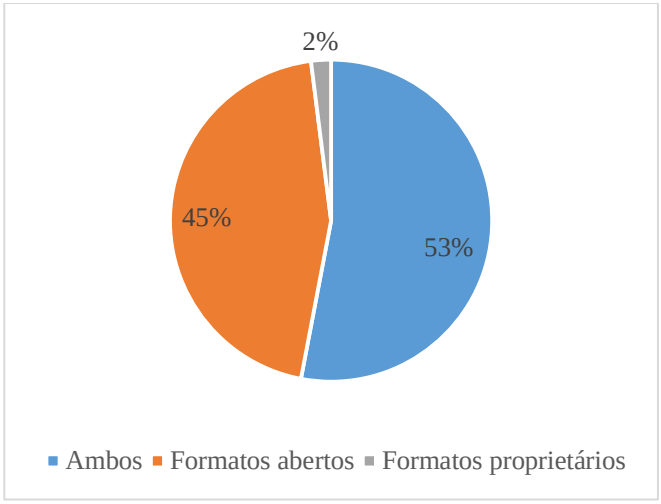

Gráfico 11 – Tipos de formato utilizados pelos RIs (N=38)

Fonte: dados da pesquisa, 2017.

Na arquitetura de referência do projeto SHAMAN (Antunes *et al.*, 2011), o conjunto de *stakeholders* é formado por 13 categorias extraídas, predominantemente, do modelo OAIS: *producer, consumer, management, executive management, information manager, technology manager, operational manager, regulator, auditor, information operator, system architect, solution provider, technology operator*.

O *framework* de política de preservação digital da *The Ohio State University Libraries* (OSUL, 2013) identificou 5 categorias de *stakeholders* extraídas e adaptadas do modelo de referência OAIS: *producer, management, administrators, co-operating archives, consumer, user groups/cliente groups*. A política deu especial atenção ao grupo

administrators, o qual foi descrito detalhadamente em uma tabela constituída por 12 elementos: *Collections Strategist, Electronic Records & Digital Resources Archivist, Head of Preservation and Reformatting, Head of Research Services, Head of the Copyright Resources Center, Head of Digital Content Services, Head of Digital Initiatives, Head, Applications Development and Support, Systems Administrator & Integration Coordinator, Curators, Metadata services staff, Technical staff.*

A responsabilidade atribuída ao *Electronic Records & Digital Resources Archivist* corresponde às funções desempenhadas pelos administradores dos repositórios pesquisados, pois eles trabalham, cooperativamente, com os setores de tecnologia da informação, com os setores de coleções especiais, com o responsável pela base de teses e de dissertações e com os responsáveis pelas comunidades e coleções. Em 66% desses RIs, a administração geral é exercida por um único bibliotecário e em 18%, por um grupo de bibliotecários. O campo outros do questionário foi preenchido por 8% dos respondentes da seguinte forma:

- "No momento o [repositório] não possui nenhum administrador formal."
- "Um bibliotecário e um analista em TI."
- "Analista de Sistemas."

Em 5% dos RIs, a administração geral é realizada por um servidor tecnoadministrativo e em 3%, por um professor (gráfico 12).

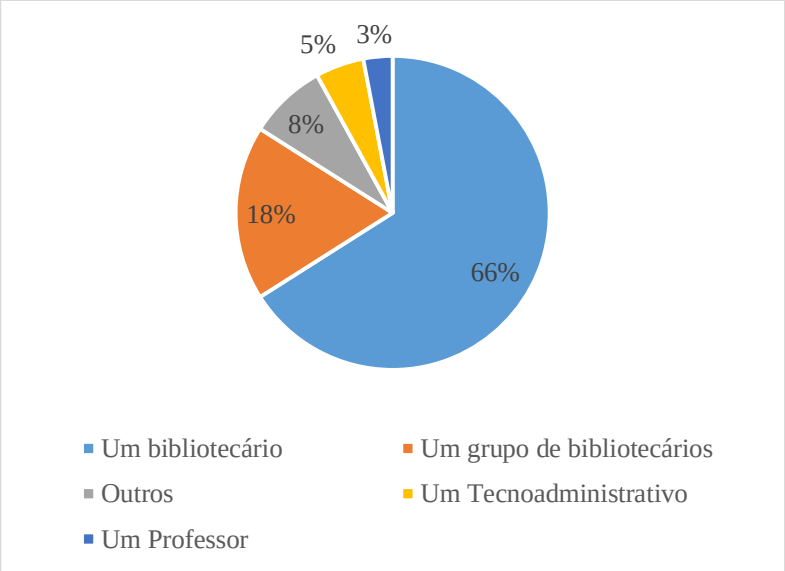

Gráfico 12 – Administrador geral dos RIs (N=38)

Fonte: dados da pesquisa, 2017.

O grupo *metadata services staff* tem como responsabilidade a de providenciar uma descrição apropriada dos metadados técnicos e administrativos para apoiar as atividades de preservação. Em 81% dos repositórios em análise, esse tipo de serviço é realizado exclusivamente por bibliotecários, em 10%, por diferenciadas funções discriminadas no campo outros:

- "O dono da coleção ou comunidade."
- "Responsável pela comunidade."
- "Servidor tecnoadministrativo (chefe da seção)."
- "01 Assistente administrativo que compõe e equipe do Sistema de Bibliotecas."

Os últimos três respondentes acima também marcaram, no questionário, a opção bibliotecário. Em 5%, o serviço é realizado por um bibliotecário e um técnico de suporte, em 4%, por um bibliotecário, um estagiário e um professor. A opção "secretário de departamento" não foi assinalada (gráfico 13).

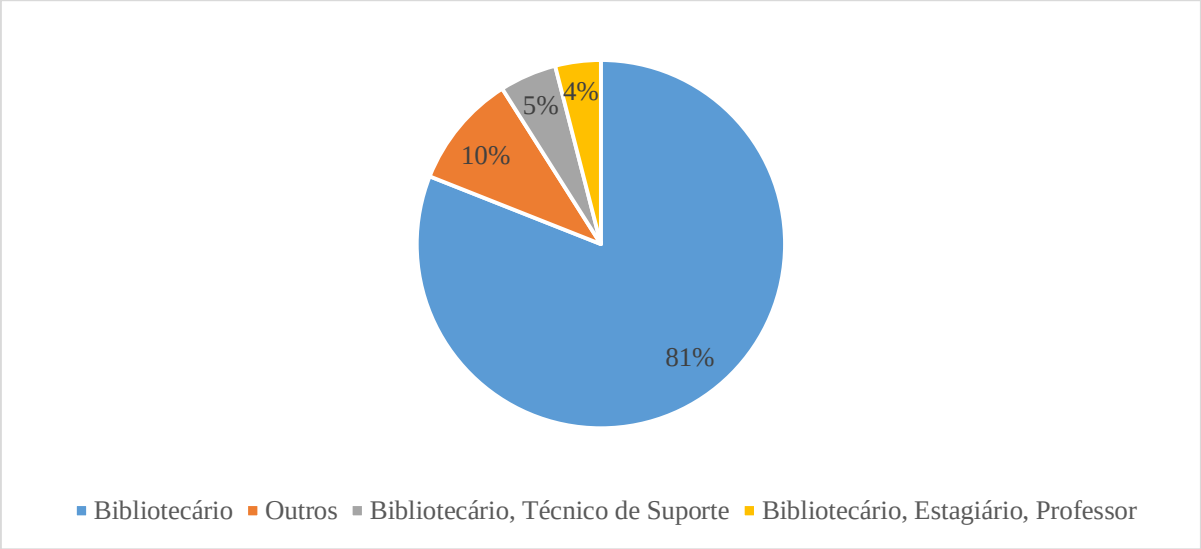

Gráfico 13 – Responsáveis pela validação de metadados nos RIs (N=38)

Fonte: dados da pesquisa, 2017.

Aos *stakeholders* da categoria *technical staff* são atribuídas as responsabilidades pelo suporte operacional das plataformas usadas para fins de preservação. Em 63% dos

RIs das universidades federais brasileiras, o suporte e a manutenção dos repositórios estão ao encargo exclusivamente dos técnicos da Superintendência (ou Núcleo) de Tecnologia da Informação; em 13%, o suporte é realizado apenas pelos técnicos de informática da Biblioteca Central; em outros, 13%, pelos técnicos discriminados no campo outros do questionário da seguinte forma:

- "Docentes e Bolsistas do [grupo de pesquisa]"
- "Técnicos do CPD"
- "servidor tecnoadministrativo (chefe da seção)"
- "Técnico de informática da Secretaria de tecnologia da Informação"
- "Assistente Administrativo"

O terceiro respondente acima também assinalou a opção "técnicos da Superintendência (ou Núcleo) de Tecnologia da Informação. Os restantes 11% são exercidos pelos técnicos de informática da Biblioteca Central (gráfico 14).

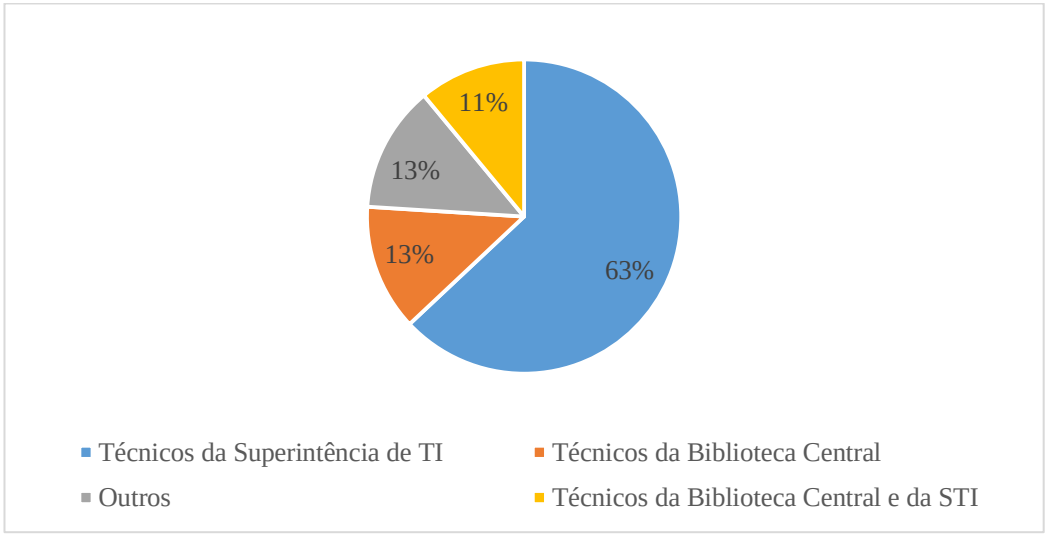

Gráfico 14 – Responsáveis pelo suporte e manutenção dos RIs (N=38)

Fonte: dados da pesquisa, 2017.

Os *stakeholders* da categoria *curators* compõem um grupo responsável pela preservação e pelo provimento do acesso aos objetos selecionados sob sua incumbência. Na seção 2.9 – Organização –, destacamos uma pesquisa realizada por Engelhardt (2013)

em dezenas de países. Este pesquisador identificou uma forte demanda de formação básica em preservação digital e planejamento de gestão de dados. Diversas iniciativas internacionais têm envidado esforços para formar curadores digitais. Por exemplo: *DigCurV Curriculum Framework*[55], *Digital Curation Center*[56], *Digital Preservation Management Workshops*[57], *Digital Preservation Training Programme*[58].

No Brasil, a Rede Brasileira de Serviços de Preservação Digital do IBICT tem parceria com 7 universidades federais, apenas: UFSM, UFRGS, UFG, UNB, UFBA, UFPB, UFRN. Em 2014, a Rede ofereceu dois cursos técnicos de preservação digital. O segundo foi totalmente dedicado à curadoria digital (IBICT, 2015). O fato de haver somente 7 universidades como uma caixa *Lockss* na rede do IBICT sinaliza um possível problema relativo à cultura de curadoria digital nesse tipo de organização pública. Porém, 50% dos respondentes afirmaram que a equipe de administradores e técnicos do repositório participou de treinamentos ou cursos de preservação digital, entre 2013 e 2016, enquanto os 50% restantes afirmaram que a equipe não participou desse tipo de atividade ao longo desses anos (gráfico 15).

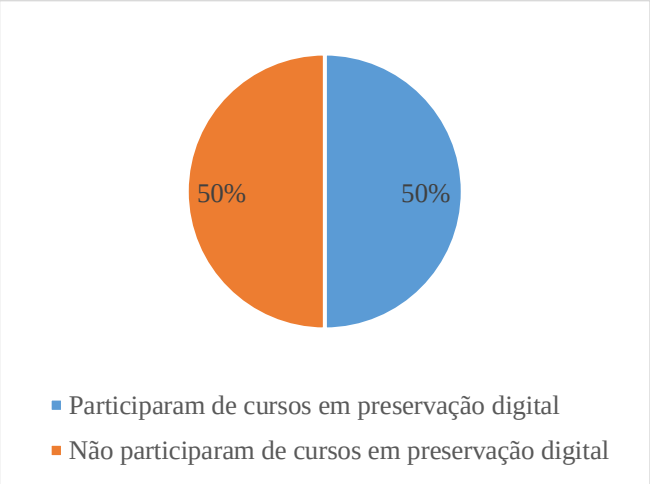

Gráfico 15 – Participação do staff dos RIs em cursos de preservação digital (N=38)

Fonte: dados da pesquisa, 2017

[55] DigCurV. Disponível na WWW: http://digcurv.gla.ac.uk/
[56] DCC. Disponível na WWW: http://www.dcc.ac.uk/
[57] DPMW. Disponível na WWW: http://www.dpworkshop.org/workshops/fiveday.html
[58] DPTP. Disponível na WWW: http://dptp.org/

Em um repositório onde se pretenda adotar uma política de procedimento de preservação de *bit*, será necessário tomar algumas providências no processo de ingestão dos objetos digitais, a fim de serem assegurados os requisitos estabelecidos pela instituição para este tipo de material. Assim, uma instituição precisa levar a cabo as seguintes medidas, minimamente: escanear os objetos com um antivírus antes da ingestão, certificar-se de que os objetos estão completos, identificar, caracterizar e validar os formatos (Sierman, Jones e ElstrØm, 2014).

No quadro 3 – Níveis de Preservação Digital –, o exame com um *antivírus* é um dos procedimentos adotados quando se desenvolvem as atividades da categoria *file fixity and data integrity* nos níveis 2 e 3. A política de preservação da OSUL (2013), ao definir o verbete *bit-level preservation* aponta a verificação de vírus como uma das formas de manter a integridade da cadeia de *bits* de um objeto digital.

O *DSpace* possui um sistema de curadoria (*Curation System*) com soluções prontas, mas é bastante flexível para ser adaptado às necessidades de um repositório específico. O objetivo desse sistema é o de facilitar o gerenciamento de taferas (*tasks*) sobre os conteúdos de um repositório. Na seção *Curation System* da documentação do sistema, foi escolhida para a exemplificação das tarefas a aplicação do *virus scan* aos *bitstreams* dos objetos armazenados. De um modo geral, as tarefas são atividades realizadas exclusivamente pelo *staff*, isto é, editores de coleção, administradores do repositório, administradores de sistema, entre outros (DURASPACE, 2015).

Nos RIs das universidades federais, 50% dos administradores não sabem informar se a base de dados dos repositórios é rastreada com o antivírus recomendado (*ClamAV*) pelo *DSpace*, 29% responderam afirmativamente quanto ao uso do antivírus e 21% responderam que os repositórios não utilizam o *ClamAV* (gráfico 16).

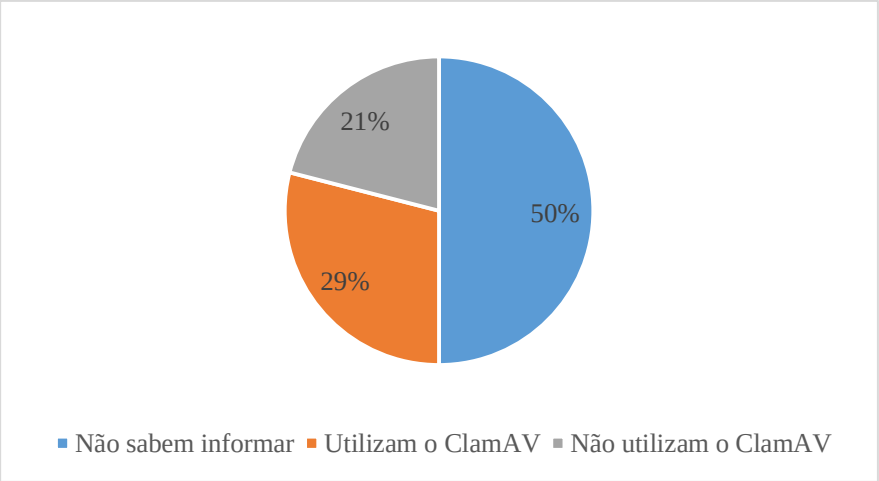

Gráfico 16 – Utilização do antivírus ClamAV (N=38)

Fonte: dados da pesquisa, 2017.

Sierman *et al.* (2014) advertem sobre o risco de se comprometer a integridade do material digital depositado nos repositórios, se um conjunto básico de atividades durante a ingestão não forem aplicadas. Dentre elas, os autores acrescentam a medida de integridade operacionalizada por um mecanismo de *fixity check*, tais como *checksums*, *message digests* and *digital signatures*. No modelo de níveis de preservação digital da NDSA (quadro 3), o *fixity check* é empregado em todos os níveis. Em outro modelo,[59] o *Digital Preservation Maturity Model* (PRESERVICA, 2014), o *file fixity* é proposto no nível 3 − *Storage Validation* − por meio do *checksum.* Um exemplo de política que assegura a integridade dos objetos digitais, tanto no processo de ingestão como em todo o ciclo de vida dos objetos pode ser lido na seção *System Security* da política de preservação digital do ICPSR (2012): "(...) ICPSR ensures the authenticity and integrity of its digital content through the active and ongoing use of checksums from receipt of the digital content onward. (...)"

A introdução da documentação do *DSpace* estimula os *stakeholders* a lerem a seção *functional overview of the system*, especialmente o pessoal que não tem função técnica. A seção tem 11 tópicos e o décimo deles − *Digital Preservation* − explica

[59] (...) In choosing to include Preservica's model [Digital Preservation Maturity Model] in the policy framework, DPTF recognized that it more suitably describes the conceptual process of digital preservation, whereas NDSA's model serves as a more effective, actionable tool for implementing digital preservation. As such, the team placed the NDSA matrix in a "parking lot" for future consideration as part of implementing a digital preservation program within OSUL. (Noonan, 2014, p.1).

detalhadamente o propósito da função *checksum checker*. Todavia, 71% dos respondentes da nossa pesquisa não souberam informar se essa funcionalidade está sendo utilizada, 16% asseguraram que a função *checksum checker* está sendo utilizada e 13% responderam que não a utilizam (gráfico 17).

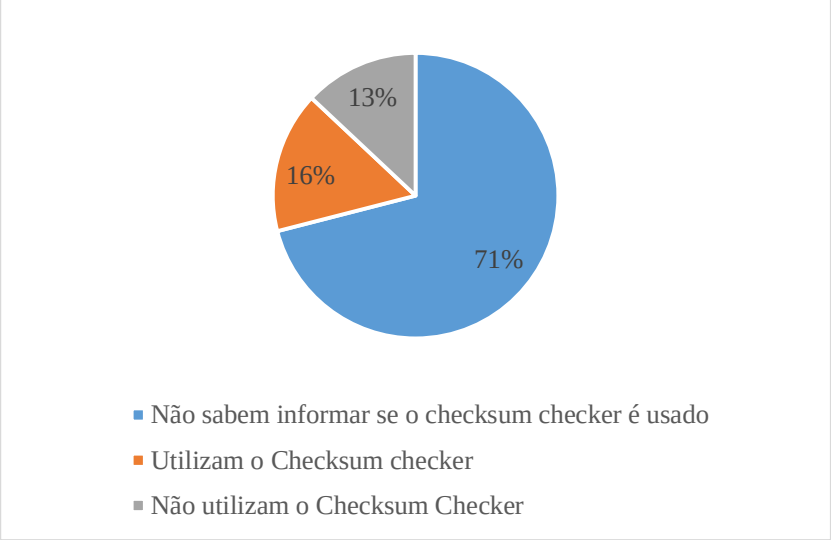

Gráfico 17 – Utilização do *checksum checker* pelos RIs (N=38)

Fonte: dados da pesquisa, 2017.

A categoria *Storage and Geographical Location* do quadro 3 recomenda, em todos os níveis de preservação, diversificar-se a localização geográfica de cópias com o fim de se proteger o material digital contra ameaças ao sistema de armazenamento, seja por causa dos desastres naturais, seja por aqueles provocados por seres humanos. No *Digital Preservation Maturity Model* (PRESERVICA, 2014), o nível 2 – *Storage Management* – sugere que o objeto digital seja mantido em pelo menos duas mídias: *hard disk* e fitas de *backup* ou em múltiplas réplicas em disco. O consórcio ICPSR adota os seguintes procedimentos para o seu sistema de armazenamento:

> ICPSR currently maintains six copies of its data (and requires that any off-site backup be encrypted):
>
> - Two copies held locally,
> - One copy put on tape once a month onsite,
> - One copy elsewhere in Michigan,
> - One copy in Amazon's storage cloud,
> - One copy with the DuraSpace cloud storage system DuraCloud (ICPSR, 2016, p. 2).

Saliente-se que o *DuraCloud* é uma das opções de *backup* para o *DSpace.* Este, por sua vez, possui duas outras opções de *backup*: *traditional backup and restore; AIP backup and restore.* A seção *AIP Backup and Restore* na documentação do *DSpace* apresenta uma tabela com as diversas diferenças entre as duas opções de *backup and restore* e sugere possíveis combinações de uso dessas opções (DURASPACE, 2015). Em 39% dos RIs pesquisados, a equipe de suporte técnico realiza apenas o *backup* tradicional; em 32%, os administradores não souberam informar se algum dos dois tipos de *backup* é realizado; em 21%, os respondentes afirmaram que os dois tipos de *backup* do *DSpace* são executados e em 8%, os respondentes afirmaram que os *backups* não são executados (gráfico 18).

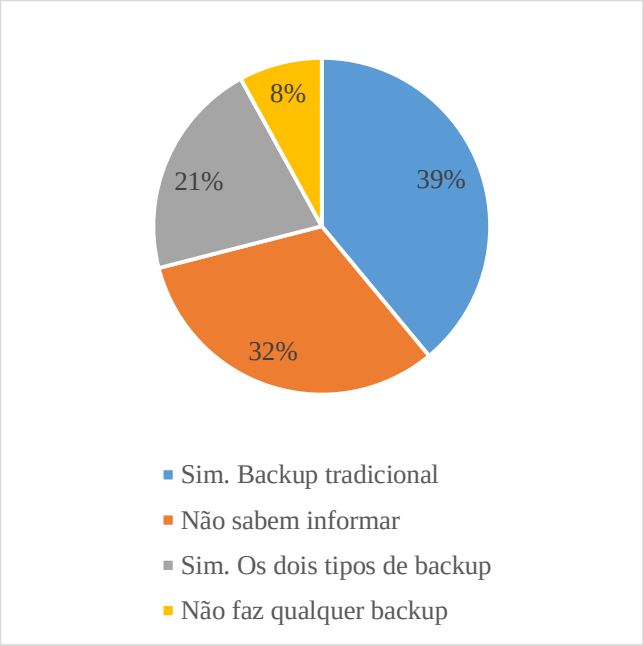

Gráfico 18 – RIs que fazem o backup tradicional e o AIP (N=38)

Fonte: dados da pesquisa, 2017.

Na seção 2.2 – Preservação ao nível do *bit* –, vimos que o uso de identificadores persistentes é considerado a melhor solução para preservar o acesso ao recurso digital, independentemente do seu URL, pois o identificador persistente será associado a uma nova localização quando o recurso for movido. Assim, uma das estratégias para se implementarem identificadores persistentes é a contratação de um sistema de

identificação persistente oferecido por terceiros, como o *Handle.Net Registry*.[60] Esse é o serviço cujo sistema é utilizado pelo *DSpace* para criar identificadores persistentes, os quais são exibidos no formato <http://hdl.handle.net/1721.123/4567>. Por outro lado, na documentação do *DSpace*, a seção *Functional Overview*, item 4.1 – *Handles* – faz a seguinte advertência:

> It is important to note that DSpace uses the CNRI Handle infrastructure only at the 'site' level. For example, in the above example [http://hdl.handle.net/1721.123/4567], the DSpace site has been assigned the prefix '1721.123'. It is still the responsibility of the DSpace site to maintain the association between a full Handle (including the '4567' local part) and the community, collection or item in question (DURASPACE, 2015, p. 4)

Em 39% dos RIs pesquisados, os administradores não sabem informar se o serviço *CNRI Handle System* foi contratado pela universidade; 37% dos respondentes afirmaram que usam o serviço e 24% afirmaram que não o utilizam (gráfico 19).

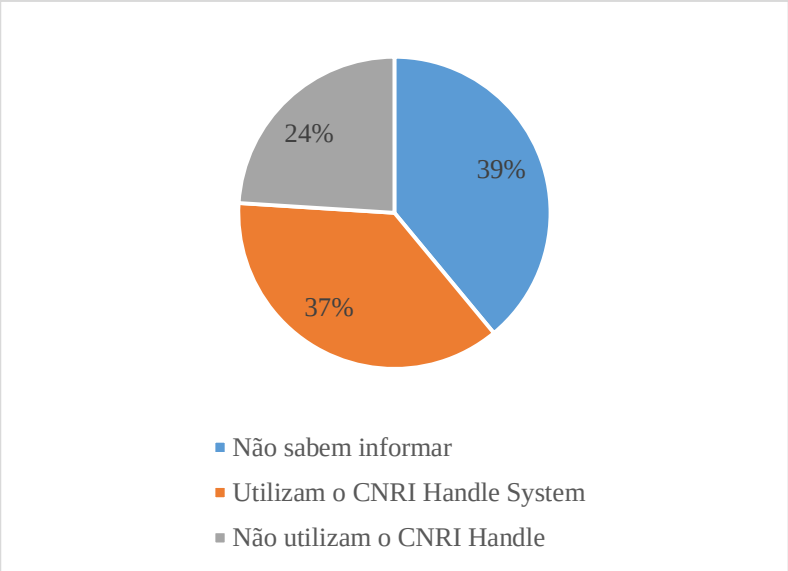

Gráfico 19 – Utilização do serviço CNRI Handle System (N=38)

Fonte: dados da pesquisa, 2017.

Na seção 2.5 – Metadados –, vimos que os metadados descritivos identificam a entidade intelectual por meio de propriedades, como autor e título, e ajudam na descoberta e acesso ao conteúdo dos objetos digitais. Esses metadados também podem determinar a

[60] Handle.Net Registry. Disponível na WWW: < http://handle.net/>

proveniência do recurso digital. Dentre os padrões adotados para estruturar essa categoria de metadados, encontram-se o *Dublin Core Metada Initiative*. Sierman *et al.* (2014) advertem que a falta de metadados descritivos podem tornar impossível a descoberta e a compreensão de um objeto digital no futuro. Um exemplo de política de preservação digital que manifesta o cuidado com a descrição do material depositado em seu repositório é o da *Wellcome Libray* (Checkley-scott e Thompson, 2014, p. 13):

> Essential to life cycle management is the generation/creation of administrative and descriptive metadata describing the material. The Library considers metadata (administrative as well as descriptive) to be essential for life cycle management and resource discovery. Metadata will be subject to similar life cycle management as the objects it describes. (...)

O *DSpace* utiliza o padrão *Dublin Core* para definir os metadados descritivos e os administrativos. Todos os 38 RIs pesquisados utilizam o *DSpace* e 71% deles possuem um manual para preenchimento de campos de metadados contra 29%, que não possuem esse tipo de *guideline* (gráfico 20).

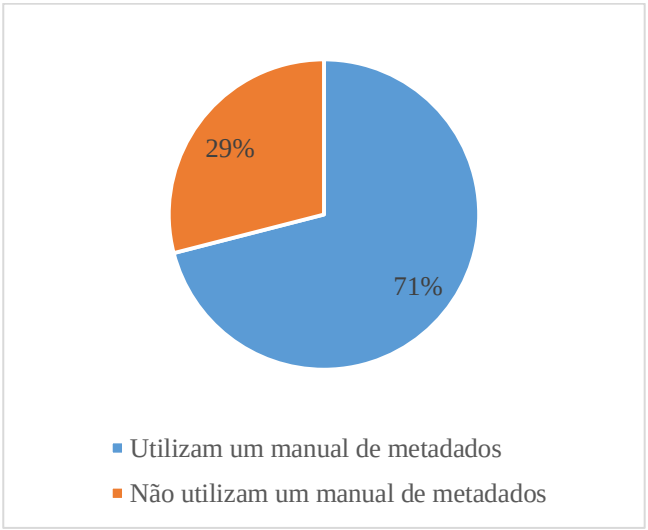

Gráfico 20 – Utilização de um manual para o preenchimento de metadados (N=38)

Fonte: dados da pesquisa, 2017.

Capítulo 6 – Proposta de um modelo de *framework* de política de preservação digital

Um *framework* de política é um conjunto de declarações explícitas que definem o nível e a natureza do compromisso e da responsabilidade institucional (Kenney *et al.*, 2014). Segundo esses autores, os sete atributos de um repositório digital confiável contêm os tipos de política, os de procedimento e outros documentos associados com cada elemento de um *framework* consolidado. Os atributos foram elaborados por Beagrie *et al*. (2002) e são resumidos assim:

> - OAIS compliance: an explicit statement that confirms the organization's commitment to complying with the Open Archival Information System (OAIS) standard.
> - Administrative responsibility: a high-level statement that demonstrates a commitment to track and comply with current and emerging standards embraced by the preservation community.
> - Organizational viability: a mission statement and comprehensive policies that document and authorize the steps an organization undertakes to receive, store, preserve, and provide access to digital materials under its care, encompassing legal, fiscal, and ethical considerations and requirements.
> - Financial sustainability: accounting and budget policies and procedures that are part of a business plan to define and protect requisite resources for the digital preservation program.
> - Technological suitability: a set of principles, policies, and procedures that define the plan for developing and maintaining requisite hardware, software, expertise, and techniques to support and enable the digital preservation program, including adherence to relevant standards and industry best practice.
> - System security: a set of policy statements and procedures that confirm the organization's commitment to maintaining a constant and appropriate level of environmental and online protection; surveillance; and risk detection, response, and mitigation to safeguard the integrity of digital assets.
> - Procedural accountability: a coherent and systematic means for documenting, sharing, and applying the set of policy statements and associated procedures and prevailing practice. These are often external to the organization itself (Kenney *et al.*, 2014, p. 6).

O guia de desenvolvimento de um *framework* de política de preservação digital elaborado por McGovern (2007) contempla todos os atributos de um repositório digital

confiável e está disponível no *website* do *Inter-university Consortium for Political and Social Research* (ICPSR, 2012), para que outras instituições[61] utilizem o guia na elaboração de suas próprias políticas. No referido guia, os atributos correspondem às seções de um *framework* com os seus respectivos componentes, acompanhados por suas descrições. Dessa forma, o *framework* é apresentado com a seguinte estrutura:

1. Compatibilidade com o modelo OAIS
2. Responsabilidade administrativa
   - Propósito
   - Objetivos
   - Mandato
3. Viabilidade organizacional
   - Escopo
   - Princípios operacionais
   - Funções e responsabilidades
   - Seleção e aquisição
   - Acesso e uso
   - Desafios e riscos
4. Sustentabilidade financeira
   - Compromisso institucional
   - Cooperação e colaboração
5. Adequação tecnológica e procedimental
6. Segurança do sistema

[61] O *framework* de preservação digital da *University of Minnesota Libraries* foi adaptado do modelo de McGovern (2007) e está disponível em www.lib.umn.edu/dp/digital-preservation-framework#8.4

7. Responsabilidade procedimental
   - Auditoria e transparência
   - Administração do *framework*
   - Definições
   - Referências

Conforme recomenda o documento *Digital Preservation Policy Tool* (ERPANET, 2003), uma política deve esclarecer como a preservação digital pode atender as principais necessidades de uma instituição e estabelecer alguns princípios e regras para sustentar um programa de implementação. No modelo do JISC (Beagrie *et al.*, 2008), a implementação (*lower level policy*) é uma seção complementar do nível mais alto (*high-level policy*), destinada para levar a efeito uma política de preservação digital promulgada. O guia de McGovern (2007) apresenta a perspectiva de alto nível de um programa de preservação digital, assim como a perspectiva de mais baixo nível (*lower level*), quando se refere às políticas e aos procedimentos de implementação. Com base nesse guia, traduzimos e adaptamos os *frameworks* de políticas de preservação digital do *Inter-university Consortium for Political and Social Research* (ICPRS, 2012) e os da *Ohio State University Libraries* (OSUL, 2013) para o modelo que passamos a propor na seção a seguir.[62] Nesse modelo, a Biblioteca Central foi escolhida como a condutora da política, porque em todas as universidades federais a Biblioteca Central é a administradora dos repositórios.

6.1 – Modelo de *framework* de política de preservação digital para as universidades federais

1) CONFORMIDADE COM O MODELO DE REFERÊNCIA OAIS

[62] A implementação do modelo poderia seguir, inicialmente, as recomendações do relatório *Master Objects Repository Task Force* (Reese *et al.*, 2014).

- Para alcançar os objetivos de preservação digital, a Biblioteca Central reconhece a necessidade de cumprir com os padrões e práticas prevalecentes da comunidade de preservação digital. A Biblioteca Central está empenhada em desenvolver suas políticas, repositórios e estratégias de preservação digital de acordo com o modelo de referência *Open Archival Information System* (OAIS, 2012).

- Se a Biblioteca Central acompanha e responde a iniciativas OAIS relacionadas, indicam-se as contribuições. Por exemplo: desenvolvimentos na certificação de arquivos digitais, de identificadores persistentes, de metadados de preservação e da interface produtor-arquivo.

- Nomear o documento que sintetiza o mapeamento do processo de preservação da Biblioteca Central para o modelo OAIS. Por exemplo: Requisitos de Preservação Digital Aplicados a Biblioteca Central.

2) RESPONSABILIDADE ADMINISTRATIVA

2.1) Propósito

- A Biblioteca Central declara o compromisso com os serviços de preservação digital e com a sustentabilidade do acesso às coleções digitais no longo prazo. Dessa forma, ficaria esclarecido quem é o principal responsável pela preservação digital.
- Na sequência da declaração anterior deve-se explicitar o alinhamento da missão da Biblioteca Central de adquirir, organizar, disseminar e preservar a produção acadêmica com o compromisso de manter um repositório digital confiável, disponível para a comunidade universitária. Dessa forma, fica esclarecido quem é o público-alvo do documento que registra a política.

- A Biblioteca Central declara que envida esforços para contribuir para a missão da universidade, explicitando as linhas gerais dessa missão. Assim, situa-se o documento no contexto mais amplo dos esforços institucionais.

2.2) Mandato

O mandato da Biblioteca Central na preservação digital poderia amparar-se nos seguintes fundamentos:

- Compromisso acadêmico — trata-se de uma declaração de que a universidade, por ser uma instituição pública de ensino superior, está comprometida com as atividades de ensino, pesquisa e extensão. À medida que mais recursos e serviços associados com essas funções forem tornando-se digitais, as responsabilidades da Biblioteca Central deverão expandir-se e incluir a identificação, gestão e preservação do conteúdo digital.

- Registros institucionais — A Biblioteca Central é responsável pela coleta e manutenção do acervo de teses e dissertações, inclusive no formato digital. Em diversas universidades, a Biblioteca Central também é responsável pelo portal de periódicos produzidos pela sua própria comunidade acadêmica.

- Obrigações legais — Neste tópico, declaram-se as responsabilidades da universidade para preservar e manter o acesso a determinados objetos digitais. Outrossim, responsabilidades são atribuídas a ela como instituição de interesse público. Algumas obrigações legais derivam de leis federais, como a Lei de Acesso à Informação [dever-se-ão acrescentar outras leis, se as houver]. Tais leis podem requerer a manutenção dos acervos em permanente atualização.

- Cumprimento de outras políticas institucionais — Esta declaração relaciona a política de preservação digital com o Estatuto da Universidade, com a Política de Informação Institucional, com a Política de Informação da Universidade, com o Plano de Desenvolvimento Institucional e com a Política de Segurança da Informação.

- Compromissos interinstitucionais — Se um desses compromissos for feito com o IBICT, dir-se-á que a universidade possui, como parceira da Rede Brasileira de Serviços de Preservação Digital (Rede Cariniana), um acordo contratual com o Instituto Brasileiro de Informação em Ciência e Tecnologia Neste sentido, a universidade é uma caixa *Lockss* que preserva periódicos científicos de acesso aberto [e o acervo dos repositórios, ou parte dele] de modo compartilhado com outros parceiros da Rede Cariniana.

2.3) Objetivos

O objetivo geral do serviço de preservação digital é o de preservar o patrimônio cultural e intelectual da universidade e assegurar que ele seja acessível e mantido de forma confiável para uso futuro. Por conseguinte, essa declaração poderia definir os objetivos específicos do *framework* nos seguintes termos:

- Identificar as coleções digitais a serem preservadas com as novas gerações de tecnologias, por meio de uma seleção sistemática.
- Manter o acesso aos objetos digitais ao nível do *bitstream* e o acesso ao significado intelectual e contextual desses objetos.
- Incluir no escopo da política os objetos nato-digitais e aqueles que foram convertidos para o formato digital.
- Proteger os investimentos da Biblioteca Central (ou outro setor da instituição que seja o responsável pela sustentabilidade financeira do programa de preservação digital) por meio de uma política de preservação digital totalmente implementada.
- Demonstrar o compromisso institucional por meio da identificação de estratégias sustentáveis.
- Desenvolver um programa eficiente com as seguintes medidas: integração de sistemas, compartilhamento de responsabilidades e automatização dos esforços humanos repetitivos.
- Cumprir as normas de preservação e acesso digital, as quais predominam na comunidade de acesso aberto.

- Buscar, expandir e desenvolver métodos de preservação digital que sejam apropriados para a universidade [denomina-se a universidade autora da política] e promover a colaboração institucional.

3) VIABILIDADE ORGANIZACIONAL

3.1) Escopo

A Biblioteca Central procede aos tópicos seguintes:

- Declara que essa política se dirige à preservação das coleções e recursos digitais, cujo agente principal de custódia é ela mesma.
- Assinala, a despeito da limitação da abrangência dessa política, que tem a responsabilidade de informar, de prestar consultoria e de cooperar com outros setores, de modo que assegura que as faculdades, os centros de ensino e a comunidade acadêmica tenham acesso adequado e contínuo aos materiais digitais produzidos na universidade fora do sistema de bibliotecas.
- Acrescentará, quando for o caso, a atuação da sua equipe (ou de todas as equipes do sistema de bibliotecas da universidade) em redes de cooperação interinstitucional, comitês normativos etc., a fim de se garantir que as faculdades, centros de ensino e a comunidade acadêmica possam acessar todos os recursos digitais de modo contínuo.
- Posiciona-se sobre a preservação de material digital que não lhe pertence ou está sob sua administração.

3.2) Princípios

Os princípios podem ser elaborados sob dois aspectos: os princípios norteadores e os princípios operacionais.

- Princípios norteadores — A Biblioteca Central adotará um critério consistente para a seleção e preservação do material digital. Uma vez que esse material

tenha sido selecionado para a administração e preservação digital, a Biblioteca Central se compromete com a manutenção pelo tempo que for necessário ou desejado, com base nos seguintes princípios:

- A Biblioteca Central está comprometida com a preservação a longo prazo do conteúdo selecionado.
- A preservação digital faz parte dos processos administrativos da Biblioteca Central.
- Os níveis de preservação e o período de tempo para manter o material digital acessível serão determinados pelos responsáveis em selecionar o material e pelos curadores digitais, com o apoio de técnicos especializados.
- A Biblioteca Central participará do desenvolvimento de padrões, práticas e soluções comunitárias de preservação digital.

- Princípios operacionais — Estes indicam os esforços que serão envidados pela Biblioteca Central:

  - Desenvolver uma infraestrutura de preservação digital escalável, confiável, sustentável e auditável.
  - Gerenciar o *hardware*, o *software* e os componentes de mídia de armazenamento, os quais possuem a função de preservação digital, conforme as normas ambientais, as especificações de controle de qualidade e os requisitos de segurança.
  - Aplicar o modelo de referência *Open Archival Information System* (OAIS) e outros padrões e práticas de preservação digital apropriados.
  - Avaliar a interoperabilidade do arquivo digital utilizando-se as opções de código aberto.
  - Assegurar a integridade dos dados.

- Estabelecer os metadados necessários para a utilização dos recursos digitais. Por exemplo: os metadados administrativos, os descritivos, os de preservação, os de proveniência, os de direitos e os técnicos.
- Respeitar os direitos do autor, os direitos de propriedade intelectual e outros direitos legais relacionados com a cópia, armazenamento, modificação e utilização dos recursos digitais.

3.3) Funções e responsabilidades

Nesta seção, declaram-se as categorias de *stakeholders* que podem executar um programa de preservação digital e são reconhecidas pela Biblioteca Central. Acrescentar-se-á que a terminologia foi adaptada do modelo de referência OAIS.

- Produtor — trata-se da função desempenhada por pessoas ou sistemas clientes que fornecem as informações a serem preservadas. Os produtores incluem professores, estudantes, funcionários, ex-alunos, colecionadores, criadores de conteúdo, editores etc. Os produtores também podem ser outros sistemas OAIS e serão responsáveis pelo cumprimento dos requisitos de depósito estabelecidos e pela gestão do arquivo digital, para assegurar uma transferência bem-sucedida (definição OAIS expandida pelo *framework* da *Ohio State University Libraries*).

- Gerenciamento — Esta é a função desempenhada por aqueles que definem a política OAIS como um componente de um domínio de políticas mais amplo. Por exemplo: a política de preservação digital será parte de uma política para toda a universidade: a Política de Informação Institucional, a Política de Desenvolvimento Institucional, a Política de Segurança da Informação, dentre outras. Um comitê formado por membros da Biblioteca Central e de outros setores envolvidos com um programa de preservação digital será o responsável

por definir as políticas de preservação digital e inseri-las em contextos organizacionais mais amplos.

- Administradores — Trata-se dos administradores de conteúdo (equipe responsável pela seleção e coleta contínua de coleções específicas), os especialistas em preservação digital e as equipes de trabalho (um apêndice com a lista de tipos de administrador com suas respectivas atribuições poderá ser acrescentado a esta política). Os administradores serão responsáveis pelo estabelecimento do programa de preservação digital e pela gestão diária dos objetos digitais.[63]

- Consumidor — Esta é a função desempenhada por pessoas ou sistemas clientes, que interagem com os serviços OAIS para encontrar informações preservadas e para acessá-las detalhadamente. Isso pode incluir pessoas ou outros sistemas com base no modelo OAIS.

- Grupos de usuários — Esta categoria discrimina os vários tipos de usuário que utilizam as coleções digitais da Biblioteca Central ou das bibliotecas setoriais.

3.4) Seleção e aquisição

Nesta categoria, relacionam-se as políticas, os formulários e os guias da Biblioteca Central, os quais orientam gestores e usuários do repositório institucional. Por exemplo:

- A Política de Desenvolvimento de Coleções estabelece as prioridades de critérios para aquisição de conteúdo digital.
- O Formulário de Depósito (*online*) reflete as prioridades e critérios estabelecidos na Política de Desenvolvimento de Coleções.

[63] Observação — No modelo OAIS, essa função é denominada de Entidade Funcional de Administração. Ela contém os serviços e as funções necessárias para controlar o funcionamento das outras entidades funcionais OAIS na rotina de um programa de preservação.

- O Guia de Preparação e Arquivamento de Dados fornece orientação e modelos para que os depositantes façam depósitos completos e bem documentados.
- A Política de Preservação Digital, no nível da implementação, estabelece os critérios e os requisitos para a preservação dos objetos digitais que serão disponibilizados a longo prazo.

3.5) Acesso e uso

Declara-se que a Biblioteca Central adquire, gerencia e preserva recursos digitais para que permaneçam acessíveis a longo prazo. Alega-se que certas limitações podem ser impostas ao acesso, devido a razões legais ou de outra ordem, mas, em geral, na medida do possível, a Biblioteca Central se esforça para tornar seus recursos digitais acessíveis a todos os usuários.

No caso de o repositório incluir a gestão de dados científicos, a Biblioteca Central declara o compromisso de proteger a identidade dos sujeitos humanos representados em dados de pesquisa, envidando esforços para desenvolver e implementar os meios para garantir a confidencialidade. Neste sentido, declaram-se as políticas de dados que existirem. Por exemplo:

- A Política de Acesso a Dados da Biblioteca Central define os princípios e critérios de acesso aos dados nas coleções do repositório institucional.
- A Política de Privacidade trata de informações sobre usuários.
- A Política de Requisição de Permissão para Redistribuir os Dados aborda o uso do conteúdo digital para outros arquivos de dados e distribuidores.

3.6) Desafios e riscos

Declara-se o reconhecimento dos desafios envolvidos na implementação de um programa de preservação digital e listam-se os que a Biblioteca Central considera ser os mais relevantes. Por exemplo:

- Crescimento acelerado do acervo — A evolução tecnológica faz surgir uma variedade de formatos de arquivo. À medida que são apresentados diferenciados tipos de material (conjunto de dados, objetos digitais complexos), torna-se premente monitorar as diferentes necessidades (tamanho do *storage*, metadados etc.) dos materiais e manter os procedimentos e as políticas que os protegem.

- Sustentabilidade — Esse desafio consiste em desenvolver-se um modelo de preservação digital sustentável que, sem subestimar ou superestimar as necessidades impostas pelas mudanças tecnológicas e de pessoal, responda a estas, conforme for necessário. A necessidade de bons modelos de custo e de programa acessíveis é amplamente reconhecida, mas ainda não totalmente abordada na instituição. A Biblioteca Central requer uma dotação orçamentária para operacionalizar e melhorar a gestão dos ativos digitais, assim como requer recursos destinados à sustentação dos esforços de preservação em curso. Acrescente-se a essa demanda financeira a existência de complexidades administrativas para assegurarem uma ação eficaz, em termos de custos e de tempo hábil, e para implementarem as estratégias de preservação. A escala de financiamento é baseada no nível de compromisso. Portanto, o programa deve refletir expectativas razoáveis de recursos necessários, ou seja, a Biblioteca Central não deve prometer mais do que pode ser entregue.

- Gestão — Passar de coleções digitais bem geridas para coleções preservadas no verdadeiro sentido do termo requer esforço institucional, desenvolvimento de parcerias e um compromisso financeiro. A Biblioteca Central deve proporcionar um senso de equilíbrio entre acesso e preservação, embora esteja consciente do papel central da preservação na manutenção do acesso.

- Parcerias — A Biblioteca Central deve trabalhar com criadores e fornecedores de conteúdo, para empregar o tratamento adequado aos objetos digitais antes do depósito. Tal medida facilitará a preservação ao longo do tempo.

- Flexibilidade — O plano de preservação digital deve ser revisto ciclicamente para responder às capacidades tecnológicas em evolução e alterar as expectativas dos usuários, sem comprometer-se o cuidado contínuo com os conteúdos digitais.

- *Expertise* — A Biblioteca Central deve comprometer-se com a formação contínua da equipe, à medida que as inovações tecnológicas forem surgindo.

- Direitos — Diversas restrições de acesso, impostas pela propriedade intelectual e por outros direitos impactam os esforços de preservação digital.

4) SUSTENTABILIDADE FINANCEIRA

A Biblioteca Central declara os recursos específicos que apoiam e aprimoram sua função na preservação digital:

4.1) Compromisso institucional — A Biblioteca Central constitui uma equipe dedicada para sustentar-lhe a função de preservação digital. Adicionalmente, o referido órgão busca recursos para investir no programa de preservação digital nas instâncias superiores da administração central da universidade. Informações detalhadas sobre estes recursos poderão ser consultadas na seção "transparência" do portal da universidade.

4.2) Cooperação e colaboração – A Biblioteca Central reconhece a preservação digital como uma responsabilidade compartilhada pela comunidade. Neste sentido, ela estabelece parcerias com outras instituições federais de ensino superior e colabora para outras instituições públicas, com as seguintes finalidades:

- Promover o desenvolvimento do programa de preservação digital.
- Partilhar lições aprendidas com outros programas de preservação digital.
- Ampliar o alcance de nossa experiência disponível.

- Estender aos usuários da Biblioteca Central, através de esforços cooperativos, o conteúdo digital que está disponível dentro de uma ampla comunidade de informação.

Em geral, no trabalho de cooperação com seus parceiros, a Biblioteca Central deseja:

- Compreender as metas, os objetivos e as necessidades das comunidades de criadores e das comunidades de consumidores de seus recursos digitais.
- Identificar parceiros e *stakeholders* a fim de contribuir para os esforços nacionais e internacionais de preservação digital.
- Ajudar a comunidade nacional e internacional a desenvolver estratégias que permitam a distribuição de atividades de coleta, descrição, digitalização, preservação e prestação de serviços.
- Trabalhar ativamente com criadores de materiais digitais para encorajar práticas e promover padrões.

5) ADEQUAÇÃO TECNOLÓGICA E DE PROCEDIMENTOS

A Biblioteca Central procede às seguintes atividades:

- Descreve o que é a maioria dos conteúdos digitais das coleções do repositório institucional.
- Lista os procedimentos a serem realizados após o recebimento de um depósito. Por exemplo: preenchimento de metadados; correção de erros; preenchimento de lacunas na documentação que acompanha o material; produção de versões para a preservação e para o acesso.
- Explica o que é feito com os arquivos (objetos digitais) que são entregues em mídias de armazenamento físico: Por exemplo: para arquivos enviados em mídia

de armazenamento físico, a Biblioteca Central faz cópias dos arquivos, mas não preserva a mídia.

- Descreve as principais estratégias de preservação digital adotadas por ela. Por exemplo: A normalização produz formatos de arquivo em ASCII (para texto) e TIFF (para imagens) a fim de se permitir a preservação e reduzir a gama de formatos de arquivo a serem preservados. Essa estratégia possibilita um melhor gerenciamento da preservação de um acervo muito heterogêneo. A migração converte o conteúdo digital em formatos de arquivo atualizados à medida que o *software* e a tecnologia relacionada vão evoluindo. Dessa forma, copia o conteúdo digital de uma mídia de armazenamento mais antiga para uma mais nova, como parte de um programa sistemático.
- Informa a sua própria iniciativa no que tange à pesquisa de outras estratégias de preservação que possam atender à crescente variedade de tipos de conteúdo digital em suas coleções.

6) SEGURANÇA DO SISTEMA

Os procedimentos de processamento dos conteúdos digitais na Biblioteca Central atentam para a necessidade de assegurar a acurácia e a integridade desses conteúdos, por meio de uma cuidadosa comparação entre os dados da documentação e os dados do conteúdo apresentado e por meio da geração de metadados.

O formulário de depósito automatizado, ao solicitar informações detalhadas e assinaturas para apresentação, procura garantir a autenticidade dos ativos digitais.

A Biblioteca Central assegura a autenticidade e a integridade dos seus conteúdos digitais, realizando *checksums* continuamente, a partir da recepção do conteúdo digital. Além disso, realiza revisões periódicas e auditorias dos seus conteúdos digitais armazenados no repositório institucional. Desenvolve vários documentos de política e procedimentos que abordam aspectos específicos da proteção a longo prazo de seus ativos digitais [referenciam-se os documentos que contêm as políticas e os procedimentos do nível de implementação].

7) RESPONSABILIDADE PROCEDIMENTAL

7.1) Auditoria e transparência

- A Biblioteca Central declara que mantém um processo contínuo de autoavaliação e melhoria que alinha políticas e práticas na biblioteca com os requisitos de Auditoria e Certificação de Repositórios Confiáveis (*Trustworthy Repository Audit & Certification – TRAC*), que foram revisados e incorporados à norma ISO / DIS 16363 (originalmente, CCSDS 652-R-1).

- Se a Biblioteca Central submeteu o repositório institucional a algum tipo de auditoria, informa-se o ano da auditoria e quem a realizou.

- A Biblioteca Central declara o compromisso com um ciclo [designando o número de anos] de autoavaliação e com um ciclo de [designando o número de anos] de auditoria para avaliar, medir e ajustar as políticas, os procedimentos, as abordagens de preservação e as práticas da função de preservação digital.

- A Biblioteca Central informa que as políticas atuais estão disponíveis em seu *website* e podem ser disponibilizadas mediante solicitação.

7.2) Administração do *framework*

A Biblioteca Central informará a data em que este *framework* de política de preservação digital foi atualizado e a data em que foi aprovado pelo Conselho Universitário. Ela se compromete em rever o *framework* designando o número de períodos anuais, para assegurar que ele permaneça atual e abrangente, à medida que a função de preservação digital na biblioteca for evoluindo.

7.3) Definições

Um glossário está disponível no apêndice [indicar a letra correspondente]. Ele fornece uma definição de cada termo usado neste *framework* de política de preservação digital.

7.4) Referências

Listam-se as fontes que foram consultadas para o desenvolvimento desse *framework.*

Considerações finais

O movimento em favor do acesso aberto ao conhecimento científico impulsionou o surgimento dos repositórios institucionais. No ano 2000, foi lançada a primeira plataforma para esse tipo de repositório, denominada de *EPrints*. O *DSpace* foi uma das principais plataformas que surgiram pouco tempo depois (Mueller, 2006; SDUM, 2015).

Um RI coleta, dissemina e preserva a produção intelectual de uma instituição. Esta, com esse dispositivo, incentiva a sua comunidade a difundir a ideologia do acesso aberto. Por conseguinte, ressaltam-se duas características fundamentais desse tipo de repositório: o acesso aberto ao conteúdo do seu acervo e ao *harvesting* dos seus metadados (Burns, Lama e Budd, 2013; SDUM, 2015; Heery e Andersen, 2005).

Assim, um RI de acesso aberto pavimenta a via verde do movimento pelo acesso aberto, valorizando o compartilhamento dos resultados das pesquisas científicas e o esforço coletivo na constituição da ciberciência. Além disso, tal dispositivo poderá elevar o papel científico, social e acadêmico de uma universidade (Borges, 2006). Várias universidades no mundo todo reconhecem a importância de constituir-se um repositório e o fazem na forma de institucionalização, por meio de instrumentos legais. Dentre elas, destacamos a Universidade do Minho (Portugal), cujo repositório é denominado de RepositóriUM. No Brasil, o IBICT encaminhou um projeto de lei para obrigar as universidades e institutos de pesquisa públicos a criarem um RI e para se formular uma política nacional de acesso livre à produção científica nacional (Kuramoto, 2009). Todavia, o projeto foi rejeitado pelo relator do processo no Senado Federal.

No estudo de Borges (2006), verificamos que o desenvolvimento de uma política institucional e as regras de utilização são peças fundamentais no processo de constituição de um RI. Destarte, o *Directory of Open Access Repositories* (OPENDOAR, 2014) alerta para a necessidade de os repositórios fornecerem informações sobre as políticas e regras de utilização. O próprio OpenDOAR elaborou uma ferramenta de políticas, assim como diversas outras iniciativas e projetos internacionais formularam modelos de política, guias e orientações diversas. Neste trabalho, exploramos alguns dos modelos de *framework* de política de preservação digital mais difundidos na literatura atual sobre esse tema.

Esta pesquisa foi realizada com os métodos quantitativo e qualitativo. Possui um alcance exploratório e descritivo, conforme conceituação dada por Sampieri, Collado e

Lucio (2013). Existem 63 universidades federais, mas apenas 38 delas possuem repositórios. Portanto, o universo (ou população) desta pesquisa é constituído por essas 38 universidades. As duas fontes de coleta de dados foram o OpenDOAR e um questionário com questões abertas e questões fechadas, baseado no *framework* do JISC, aplicado aos administradores dos repositórios. A análise empreendida nesta pesquisa procurou, primeiramente, verificar se alguma política de preservação digital fora derivada das PIIs encontradas nas *homepages* dos repositórios. Em seguida, após a aplicação do questionário, analisamos os sentidos dos dados colhidos e quantificados; relacionamos as descobertas com o referencial teórico abordado.

As universidades federais brasileiras criaram seus RIs a partir de uma iniciativa do IBICT, o qual contemplou dezenas dessas instituições com computadores equipados com o *software DSpace*. Em contrapartida, as universidades deveriam criar uma PII, cujo modelo de f*ramework* foi sugerido pelo IBICT (Kuramoto, 2009; Ribeiro, 2012).

O levantamento dos repositórios federais foi iniciado mediante o exame do portal *OpenDOAR*. Neste, descobrimos 26 repositórios cadastrados. Além disso, nos chamou a atenção o fato de a ferramenta de políticas não estar preenchida em todos os repositórios cadastrados, especialmente a ferramenta de política de preservação digital. Dentre estes repositórios, apenas 16 possuem uma PII e, em todas elas, manifesta-se a preocupação com a preservação digital, embora não exista uma PPD criada pela instituição. Esta descoberta corrobora os estudos de autores como Ribeiro (2012) e Medeiros e Ferreira (2014), os quais concluem que a cultura de preservação digital nas instituições federais de ensino superior ainda é escassa.

As universidades que cadastraram seus respectivos RIs no DOAR fizeram um esforço de aumentar a visibilidade desses repositórios, mas não utilizaram a ferramenta de políticas do diretório. Uma situação semelhante pode ser verificada pelo exame das PIIs, nas quais as intenções de preservação não se concretizaram por meio de uma PPD.

Assim, para a universidade levar a cabo o interesse de preservar o acervo do repositório por meio de um programa, uma política de preservação digital deve ser aprovada pelo Conselho Universitário a fim de ser oficializado o compromisso institucional com as ações relativas à preservação digital. Uma política bem estruturada depende do investimento em capacitação do *staff* do repositório. Neste sentido, as

universidades poderiam estabelecer parcerias educacionais com o IBICT para este treinar o pessoal envolvido no planejamento e nas práticas da preservação digital.

A análise documental evidenciou a consciência das universidades federais sobre a importância da preservação digital, mas não foi suficiente para revelar a existência (ou a inexistência) de práticas de preservação nessas instituições nem a existência dos demais elementos que podem estruturar a elaboração de uma política de preservação digital. Sendo assim, um outro levantamento (uso de um questionário) se fez necessário para explorarmos a cultura de preservação digital em todas as universidades detentoras de um RI.

O questionário procurou identificar alguns elementos necessários para se formular o nível superior e o inferior de uma política de preservação digital, com base nos critérios do modelo de *framework* do JISC (Beagrie *et al.*, 2008). Assim, na primeira parte do questionário, notamos a ausência de alguns elementos na maioria dos repositórios. Por exemplo: os administradores de comunidades não reclamam contra a falta de uma PPD; os repositórios não estabelecem um acordo para realizar a conversão de formato de arquivo para fins de preservação; os repositórios não possuem um glossário de termos relacionados com a preservação digital. Na segunda parte do questionário, também identificamos a ausência de alguns elementos necessários para se implementar uma política, na maioria dos repositórios. Por exemplo: a inexistência de uma cláusula de preservação nas licenças aplicadas; a desinformação sobre o uso do antivírus ClamAV; a desinformação sobre a utilização da ferramenta *checksum checker*.

Nesta pesquisa, alcançamos os objetivos específicos da seguinte forma: no capítulo 1, analisamos as grandes mudanças na academia com a emergência da ciberciência e caracterizamos a origem e a função dos RIs. No capítulo 3, analisamos os modelos de *framework* de política de preservação digital elaborados com iniciativas europeias e norte-americanas. No capítulo 5, primeiramente, verificamos se existia uma política de preservação digital para os RIs das universidades federais brasileiras e, em seguida, apuramos a percepção dos administradores desses repositórios no tocante às questões relativas à preservação digital. Quanto ao objetivo geral, este foi alcançado no capítulo 6, onde propusemos um modelo de *framework* de política de preservação digital para os repositórios institucionais das universidades federais brasileiras.

A nossa questão de pesquisa indagou das iniciativas adotadas pelas universidades federais do Brasil. Tais iniciativas poderiam contribuir para a definição de um modelo de política de preservação digital em seus repositórios institucionais. Como resposta descobrimos que pouco mais da metade das universidades pesquisadas possui um setor responsável pela preservação digital. A maioria delas pretende ingressar o seu RI na Rede Brasileira de Serviço de Preservação Digital, coordenada pelo IBICT. Em metade dessas universidades, as equipes de técnicos e administradores participaram de treinamentos em preservação digital nos últimos dois anos. A maioria dos RIs realiza os *backups* recomendados pelo *DSpace* e possui um manual para o preenchimento de metadados.

As universidades federais dão grande importância ao acesso às informações administrativas. Para tanto, disponibilizam seções em seus portais, dedicadas à transparência, e ligam os portais ao sistema de informação do Governo Federal para demonstrarem que estão conscientes de suas obrigações com a Lei de Acesso à Informação etc. Entretanto, aquelas que criaram um RI precisam ampliar suas ações de preservação digital, de modo que garantam o acesso à sua produção acadêmica e científica, a longo prazo. Para tanto, o passo inicial é dado com a criação de uma política de preservação digital.

Com este trabalho, pretendemos contribuir para a formulação de uma política de preservação digital, no Brasil, tomando por base as melhores práticas internacionais e tendo como alvos de aplicação as universidades federais. Nestas instituições, pudemos comprovar a completa ausência desse tipo de política, apesar da crescente participação da academia brasileira no movimento em prol do acesso aberto[64] e do aumento do acervo dos repositórios pesquisados. Esses fatos justificam plenamente a formulação e a implementação desse tipo de política com a participação de todos os *stakeholders*. Cabe-nos advertir que, sem o envolvimento da comunidade acadêmica e das instâncias superiores da administração universitária, não será possível avançar eficazmente na preservação da produção intelectual da instituição. Consequentemente, o acesso a essa produção ficará comprometido, pois a cada dia essa produção é, preponderantemente, digital.

[64] Visível, entre outras participações, na realização conjunta de uma conferência dedicada à sua discussão no Brasil e em Portugal, a CONFOA.

Referências

ADAM, Sharon - Preserving authenticity in the digital age. **Library Hi Tech**. [Em linha] 28:4 (2010) 595–604. [Consult. 3 nov. 2014]. Disponível em WWW:<URL:http://www.emeraldinsight.com/doi/10.1108/07378831011096259>. ISSN 0737-8831.

ANTUNES, Gonçalo *et al.* - **Shaman Reference Architecture** [Em linha]. [S.l.] : SHAMAN, 2011 (Relatório n.3). [Consult. 5 out. 2016]. Disponível em WWW:<URL:http://algos.inesc-id.pt/~jpa/InscI/poisson/varwwwhtml/portal/ficheiros/publicacoes/8263.pdf>.

ARCHIVES NEW ZEALAND - **Glossary Digital Continuity Definitions** [Em linha]. Wellington : Department of Internal Affairs, 2015, atual. 2015. [Consult. 12 jun. 2015]. Disponível em WWW:<URL:http://archives.govt.nz/advice/continuum/glossary/digital-continuity-definitions>.

ATKINS, Winston *et al.* - **Staffing for Effective Digital Preservation** [Em linha]. Washington, D. C. : National Stewardship Alliance, 2013, atual. 2013. [Consult. 14 mai. 2015]. Disponível em WWW:<URL:l: http://hdl.loc.gov/loc.gdc/lcpub.2013655113.1>.

BEAGRIE, Neil *et al.* - **Trusted Digital Repositories: Attributes and Responsibilities** [Em linha]. Mountain View : Research Libraries Group, 2002 [Consult. 3 out. 2016]. Disponível em WWW:<URL:http://www.oclc.org/research/activities/past/rlg/trustedrep/repositories.pdf>.

BEAGRIE, Neil *et al.* - **Digital Preservation Policies Study: Part 1 - Final Report** [Em linha]. Salisbury : Charles Beagrie Limited, 2008 (Relatório n.1). [Consult. 12 jan. 2017]. Disponível em WWW:<URL:http://www.jisc.ac.uk/media/documents/programmes/preservation/jiscpolicy_p1finalreport.pdf>.

BEAGRIE, Neil; JONES, Maggie - **Preservation Management of Digital Materials : The Handbook** [Em linha]. Heslington : Digital Preservation Coalition, 2008 [Consult. 8 jun. 2015]. Disponível em WWW:<URL:http://www.dpconline.org/component/docman/doc_download/299-

digital-preservation-handbook>.

BECKER, Christoph *et al.* - Systematic planning for digital preservation. **Int. J. Digit. Libr.** [Em linha] 10:4 (2009) 133–157. [Consult. 9 fev. 2015]. Disponível em WWW:<URL:http://link.springer.com/article/10.1007/s00799-009-0057-1>. ISSN 1432-1300.

BELLINI, Emanuele *et al.* - **Persistent Identifiers Ineroperability Framework** [Em linha]. Didcot : APARSEN, 2012 (Relatório n.D22.1). [Consult. 12 jul. 2016]. Disponível em WWW:<URL:http://www.alliancepermanentaccess.org/wp-content/plugins/download-monitor/download.php?id=81>.

BESEK, June M. *et al.* - Digital Preservation and Copyright: An International Study. **International Journal of Digital Curation**. [Em linha] 3:2 (2008) 103–111. [Consult. 18 jan. 2017]. Disponível em WWW:<URL:http://www.ijdc.net/index.php/ijdc/article/view/90>. ISSN 1746-8256.

BOAI - **Budapest Open Access Initiative** [Em linha], atual. 2002. [Consult. 20 mai. 2016]. Disponível em WWW:<URL:http://www.budapestopenaccessinitiative.org/boai-10-recommendations>.

BORGES, Maria Manuel - **A Esfera - Comunicação Acadêmica e Novos Media**. Coimbra : Universidade de Coimbra, 2006 Tese.

BROWN, Adrian - **Practical Digital Preservation: a how-to guide for organizations of any size**. 1. ed. London : Facet Publishing, 2013. ISBN 9781856049627.

BURNS, C.Sean; LANA, Amy; BUDD, John M. - Institutional repositories: Exploration of costs and value. **D-Lib Magazine**. [Em linha] 19:1–2 (2013). [Consult. 4 mai. 2015]. Disponível em WWW:<URL:http://www.dlib.org/dlib/january13/burns/01burns.html>. ISSN 10829873.

CANDELA, L. *et al.* - **The Digital Library Reference Model** [Em linha]. Glasgow : DL.org, 2011 (Relatório n.231551). [Consult. 2 jun. 2015]. Disponível em WWW:<URL:http://bscw.research-infrastructures.eu/pub/bscw.cgi/d222816/D3.2b Digital Library Reference Model.pdf>.

CAPLAN, Priscilla - **Metadata fundamentals for all librarians** [Em linha].

Washington, USA : American Library Association, 2003 [Consult. 2 mar. 2015]. Disponível em WWW:<URL:https://books.google.pt/books?id=yt2863FismcC&lpg=PP1&hl=pt-BR&pg=PR2#v=onepage&q&f=false>. ISBN 0838908470.

CARVALHO, José *et al.* - Auditoria ISO 16363 a repositórios institucionais. **Cadernos BAD**. Coimbra. [Em linha]2 (2014) 29–39. [Consult. 13 mai. 2013]. Disponível em WWW:<URL:http://www.bad.pt/publicacoes/index.php/cadernos/article/view/1175>. ISSN 1645-2895.

CASTELLS, Manuel - **A era da informação: economia, sociedade e cultura**. 6. ed. São Paulo : Paz e Terra, 1999. ISBN 9788577530366.

CCSDS - CONSULTATIVE COMMITTEE FOR SPACE DATA SYSTEMS - **Reference Model for an Open Archival Information System (OAIS)** [Em linha]. Washington, DC : CCSDS Secretariat, 2012 [Consult. 14 mai. 2015]. Disponível em WWW:<URL:http://public.ccsds.org/publications/archive/650x0b1.pdf>.

CDL - CALIFORNIA DIGITAL LIBRARY - **Unified Digital Format Registry** [Em linha]. San Francisco : Univeristy of California, 2012 [Consult. 15 mai. 2015]. Disponível em WWW:<URL:http://udfr.org/project/UDFR-final-report.pdf>.

CHAUI, Marilena - A universidade pública sob nova perspectiva. **Revista Brasileira de Educação**. . ISSN 1413-2478. 24 (2003) 5–15. doi: 10.1590/S1413-24782003000300002.

CHECKLEY-SCOTT, Caroline; THOMPSON, Dave - **Wellcome Library Preservation Policy for Materials Held in Collections** [Em linha]. [S.l.] : Wellcome Library, 2014 (Relatório n.2). [Consult. 6 out. 2016]. Disponível em WWW:<URL:http://wellcomelibrary.org/content/documents/policy-documents/preservation-policy>.

CHIN - **Intellectual Property Rights** [Em linha]. [S.l.] : Canada.ca, 2016, atual. 2016. [Consult. 3 set. 2016]. Disponível em WWW:<URL:http://canada.pch.gc.ca/eng/1445528228222/1445528228225>.

COLEY, Karen - Descriptive metadata for copyright status. **First Monday**. Illinois. [Em linha] 10:10 (2005) 1–9. [Consult. 17 mai. 2015]. Disponível em

WWW:<URL:http://journals.uic.edu/ojs/index.php/fm/article/view/1282>. ISSN 13960466.

COUTINHO; PEREIRA, Clara Maria Gil Fernandes - **Metodologia de Investigação em Ciências Sociais e Humanas: Teoria e Prática**. 2. ed. Coimbra : Edições Almedina, 2015. ISBN 9789724051376.

CTDE - CÂMARA TÉCNICA DE DOCUMENTOS ELETRÔNICOS - **eARQ Brasil: modelo de requisitos para sistemas informatizados de gestão arquivística de documentos** [Em linha]. Rio de Janeiro, Brasil : CONARQ, 2011 [Consult. 5 mar. 2015]. Disponível em WWW:<URL:http://www.conarq.arquivonacional.gov.br/media/publicacoes/earq/conarq_earqbrasil_model_requisitos_2009.pdf>. ISBN 978-85-60207-30-5.

CTDE - CÂMARA TÉCNICA DE DOCUMENTOS ELETRÔNICOS - **Glossário - Documentos Arquivísticos Digitais** [Em linha]. 6. ed. Rio de Janeiro : CONARQ, 2014 [Consult. 12 jun. 2015]. Disponível em WWW:<URL:http://www.documentoseletronicos.arquivonacional.gov.br/media/publicacoes/glossario/2014ctdeglossario_v6_public.pdf>.

CTDE - CÂMARA TÉCNICA DE DOCUMENTOS ELETRÔNICOS - **Diretrizes para a implementação de repositorios digitais confiáveis de documentos arquivísticos** [Em linha]. Rio de Janeiro : CONARQ, 2014 [Consult. 17 mai. 2015]. Disponível em WWW:<URL:http://www.documentoseletronicos.arquivonacional.gov.br/media/publicacoes/repositorios/conarq_repositorios_completa.pdf>.

DAPPERT, Angela; ENDERS, Markus - Digital Preservation Metadata Standards. **Information Standards Quarterly**. Baltimore. [Em linha] 22:2 (2010) 4–13. [Consult. 14 mai. 2015]. Disponível em WWW:<URL:http://www.niso.org/apps/group_public/download.php/4236/FE_Dappert_Enders_MetadataStds_isqv22no2.pdf>. ISSN 1041-0031.

DAY, Michael; ROSS, Seamus (Eds) - Metadata. **DCC Digital Curation Manual**. Edinburgh. [Em linha] 1:1 (2005) 1–41. [Consult. 18 mai. 2015]. Disponível em WWW:<URL:http://www.dcc.ac.uk/sites/default/files/documents/resource/curation-manual/chapters/metadata/metadata.pdf>. ISSN 1747-1524.

DCC - DIGITAL CURATION CENTER - **Curation Lifecycle Model** [Em linha]. Edinburgh : DCC, 2015, atual. 2015. [Consult. 2 jun. 2015]. Disponível em WWW:<URL:http://www.dcc.ac.uk/resources/curation-lifecycle-model>.

DIAS, Guilherme Ataíde; SOUSA, Rosilene Paiva Marinho; PAIVA, Maria Jose Rodrigues - Direito autoral e preservação digital : considerações pertinentes a periódicos científicos eletrônicos mantidos no sistema LOCKSS. **Ci. Inf.** Brasília. [Em linha] 41:1 (2012) 92–102. [Consult. 6 abr. 2015]. Disponível em WWW:<URL:http://revista.ibict.br/ciinf/index.php/ciinf/article/view/2117>. ISSN 1518-8353.

DIGCURV - **A Curriculum Framework for Digital Curation** [Em linha]. [S.l.] : Digital Curator Vocational Education Europe Project, 2013 [Consult. 14 mai. 2015]. Disponível em WWW:<URL:http://www.digcurv.gla.ac.uk/index.html>.

DPH - **Digital Preservation Handbook** [Em linha]. Glasgow : Digital Preservation Coalition, 2015, atual. 2015. [Consult. 4 out. 2016]. Disponível em WWW:<URL:http://handbook.dpconline.org/>.

DSA-DATA SEAL OF APPROVAL - **About** [Em linha] [Consult. 7 jul. 2016]. Disponível em WWW:<URL:http://www.datasealofapproval.org/en/information/about/>.

DSPACE@MIT - **About DSpace@MIT: Home** [Em linha]. Cambridge : MIT Libraries, 2016, atual. 2016. [Consult. 27 jul. 2016]. Disponível em WWW:<URL:http://libguides.mit.edu/c.php?g=176372&p=1158829>.

DURANTI, Luciana - Diplomatics : New Uses for an Old Science. **Archivaria**. [Em linha]28 (1989) 7–27. [Consult. 4 dez. 2014]. Disponível em WWW:<URL:http://journals.sfu.ca/archivar/index.php/archivaria/article/viewFile/11567/12513>. ISSN 1923-6409.

DURASPACE - **Curation System** [Em linha]. [S.l.] : DSpace, 2015, atual. 2015. [Consult. 6 out. 2016]. Disponível em WWW:<URL:https://wiki.duraspace.org/display/DSDOC5x/Curation+System#CurationSystem-VirusScan>.

DURASPACE - **AIP Backup and Restore** [Em linha]. [S.l.] : DSpace, 2015, atual.

2015. [Consult. 5 out. 2016]. Disponível em WWW:<URL:https://wiki.duraspace.org/display/DSDOC5x/AIP+Backup+and+Restore>.

DURASPACE - **Persistent URLs and Identifiers** [Em linha]. [S.l.] : DSpace, 2015, atual. 2015. [Consult. 6 out. 2016]. Disponível em WWW:<URL:https://wiki.duraspace.org/display/DSDOC5x/Functional+Overview#FunctionalOverview-PersistentURLsandIdentifiers>.

ENGELHARDT, Claudia - The DigCurV Review of Training Needs in the Field of Digital Preservation and Curation. Em **Framing the Digital Curation Curriculum Conference** [Em linha]. Florence : Digital Curator Vocational Education Europe Project, 2013 [Consult. 14 mai. 2015]. Disponível em WWW:<URL:http://www.digcur-education.org/eng/Resources/DigCurV-2013-proceedings/Engelhardt-paper10>.

ERPANET - **Digital Preservation Policy Tool** [Em linha]. Glasgow : [s.n.] [Consult. 12 jan. 2017]. Disponível em WWW:<URL:http://www.erpanet.org/guidance/index.php#policy>.

FERREIRA, Miguel *et al.* - Carrots and Sticks: Some Ideas on How to Create a Successful Institutional Repository. **D-Lib Magazine**. [Em linha] 14:1/2 (2008). [Consult. 1 jul. 2014]. Disponível em WWW:<URL:http://www.dlib.org/dlib/january08/ferreira/01ferreira.html>. ISSN 1082-9873.

FERREIRA, Miguel; SARAIVA, Ricardo; RODRIGUES, Eloy - **Estado da Arte em Preservação Digital** [Em linha]. [S.l.] : RCAAP, 2012 [Consult. 19 jul. 2016]. Disponível em WWW:<URL:http://hdl.handle.net/1822/17049>.

FURG - **Política de Informação Institucional da FURG** [Em linha], atual. 2011. [Consult. 13 mai. 2016]. Disponível em WWW:<URL:http://repositorio.furg.br/static/politica>.

GARRETT, J.; WATERS, D. - **Preserving Digital Information** [Em linha]. Washington, DC : CLIR, 1996 [Consult. 18 mai. 2015]. Disponível em WWW:<URL:http://www.clir.org/pubs/reports/pub63>.

GARTNER, Richard - **Metadata for digital libraries: state of the art and future**

directions [Em linha]. Bristol : JISC, 2008 [Consult. 1 mar. 2015]. Disponível em WWW:<URL:http://www.jisc.ac.uk/media/documents/techwatch/tsw_0801pdf.pdf>.

GARTNER, Richard; LAVOIE, Brian - Preservation Metadata. **DPC Technology Watch Series**. [Em linha] 1:1 (2013) 1–15. [Consult. 3 mar. 2015]. Disponível em WWW:<URL:http://www.dpconline.org/component/docman/doc_download/894-dpctw13-03>. ISSN 20487916.

GIARRETA, David - **Advanced Digital Preservation**. ISBN 978-3-642-16808-6.

GIARRETA, David; HARMSEN, Henk; KEITEL, Christian - **Memorandum of Understanding** [Em linha]. [S.l.] : Trusted Digital Repository, 2010 [Consult. 13 mai. 2015]. Disponível em WWW:<URL:http://www.trusteddigitalrepository.eu/Memorandum of Understanding.html>.

GLADNEY, Henry M. - **Preserving Digital Information**. Heidelberg : Springer-Verlag, 2007. ISBN 978-3-540-37886-0.

GRACE, Stephen; KNIGHT, Gareth; MONTAGUE, Lynne - **InSPECT: Final Report** [Em linha]. London : InSPECT Project, 2009 (Relatório n.1). [Consult. 18 mai. 2015]. Disponível em WWW:<URL:http://www.significantproperties.org.uk/inspect-finalreport.pdf>.

HARNAD, Steven - Scholarly Skywriting and the Prepublication Continuum Inquiry. **Psychological Science**. [Em linha]1 (1990) 342–343. [Consult. 19 jun. 2014]. Disponível em WWW:<URL:http://cogprints.org/1581/1/harnad90.skywriting.html>.

HARNAD, Steven - Post-Gutenberg Galaxy: The Fourth Revolution in the Means of Production of Knowledge. **The Public - Access Computer Systems Review**. [Em linha] 2:1 (1991) 39–53. [Consult. 11 jun. 2014]. Disponível em WWW:<URL:https://journals.tdl.org/pacsr/index.php/pacsr/article/view/6030/5662>.

HEDSTROM, Margaret - Digital Preservation: A Time Bomb for Digital Libraries. **Computers and the Humanities**. [Em linha] 31:1998) 189–202. [Consult. 18 mai. 2015]. Disponível em WWW:<URL:http://deepblue.lib.umich.edu/bitstream/handle/2027.42/42573/?sequence=1>. ISSN 0010-4817.

HEERY, Rachel; ANDERSON, Sheila - **Digital Repositories Review** [Em linha]. Bath : JISC, 2005 [Consult. 20 mai. 2015]. Disponível em WWW:<URL:http://opus.bath.ac.uk/23566/2/digital-repositories-review-2005.pdf>.

HOUGHTON, Bernadette - Trustworthiness: Self-assessment of an Institutional Repository against ISO 16363-2012. **D-Lib Magazine**. [Em linha] 21:3/4 (2015) atual. Mar./Ab. 2015. [Consult. 13 mai. 2015]. Disponível em WWW:<URL:http://www.dlib.org/dlib/march15/houghton/03houghton.html>. ISSN 1082-9873.

IBICT - **Manifesto Brasileiro de Apoio ao Acesso Livre à Informação Científica** [Em linha]. Brasília : [s.n.], atual. 2005. [Consult. 4 out. 2016]. Disponível em WWW:<URL:http://livroaberto.ibict.br/Documentos.jsp>.

IBICT - **Rede Cariniana** [Em linha]. Brasília : Instituto Brasileiro de Informação em Ciência e Tecnologia, 2015, atual. 2015. [Consult. 4 out. 2016]. Disponível em WWW:<URL:http://cariniana.ibict.br/index.php/inicio>.

ICPSR - **Digital Preservation Policy Framework** [Em linha]. Ann Arbor : The Inter-university Consortium for Political and Social Research, 2012, atual. 2012. [Consult. 4 out. 2016]. Disponível em WWW:<URL:http://www.icpsr.umich.edu/icpsrweb/content/datamanagement/preservation/policies/dpp-framework.htm>.

ICPSR - **Archival Storage** [Em linha]. Ann Arbor : [s.n.], atual. 2016. [Consult. 5 out. 2016]. Disponível em WWW:<URL:http://www.icpsr.umich.edu/icpsrweb/content/datamanagement/preservation/storage.html>.

INTERPARES - INTERNACIONAL RESEARCH ON PERMANENT AUTHENTIC RECORDS IN ELETRONIC SYSTEMS - **Chain of Preservation Model** [Em linha]. Vancouver : InterPARES, 2007 [Consult. 2 jun. 2015]. Disponível em WWW:<URL:http://www.interpares.org/ip2/ip2_model_display.cfm?model=cop>.

INTERPARES (INTERNACIONAL RESEARCH ON PERMANENT AUTHENTIC RECORDS IN ELETRONIC SYSTEMS) - **Module2: Developing Policy and Procedures for Digital Preservation** [Em linha]. Vancouver : InterPARES, 2012

[Consult. 27 jul. 2016]. Disponível em WWW:<URL:http://www.interpares.org/ip3/display_file.cfm?doc=ip3_canada_gs12_module_2_july-2012_DRAFT.pdf>.

ISO 16363. 2012, Space Data And Information Systems - **Audit and certification of trustworthy digital repositories**. 1. ed. Geneva : CCSDS Secretariat, 2012

ISO 16919. 2014, Space Data And Information Systems - **Requirements for Bodies Providing Audit and Certification of Candidate Trustworthy Digital Repositories**. 1. ed. Geneva : CCSDS Secretariat, 2014

ISO 20562. 2006, Space Data And Information Transfer Systems - **Producer-archive interface -- Methodology abstract standard**. 1. ed. Geneva : CCSDS Secretariat, 2006

JISC INFONET - **Digital Repositories infoKit** [Em linha]. Newcastle : University of Northumbria, 2012 [Consult. 20 mai. 2015]. Disponível em WWW:<URL:http://tools.jiscinfonet.ac.uk/downloads/repositories/digital-repositories.pdf>.

JWML - J. WILLARD MARRIOTT LIBRARY - **Digital Preservation Policy** [Em linha]. Salt Lake City : The University of Utah, 2012 [Consult. 14 mai. 2015]. Disponível em WWW:<URL:http://www.lib.utah.edu/collections/digital/DigitalPreservationPolicy2012.docx>.

KAUR, Kirnn *et al.* - **Report on DRM Preservation** [Em linha]. Didcot : APARSEN, 2014 (Relatório n.D31.1). [Consult. 14 mai. 2015]. Disponível em WWW:<URL:http://www.alliancepermanentaccess.org/wp-content/uploads/downloads/2014/06/APARSEN-REP-D31_1-01-1_4_incURN.pdf>.

KEITEL, Christian - **DIN Standard 31644 and Nestor Certification** [Em linha]. [S.l.] : Fondazione Rinascimento Digitale, 2012, atual. 2012. [Consult. 13 mai. 2015]. Disponível em WWW:<URL:http://93.63.166.138:8080/dspace/handle/2012/99>.

KENNEY, Anne R. *et al.* - **Digital Preservation Management: Implementing Short-Term Strategies for Long-Term Problems** [Em linha]. Massachusetts : MIT Libraries, 2014 [Consult. 5 out. 2016]. Disponível em WWW:<URL:http://www.dpworkshop.org/dpm-eng/eng_index.html>.

KING, Ross - **SCAPE Project Ends on the 30th of September** [Em linha], atual. 2014. [Consult. 23 set. 2014]. Disponível em WWW:<URL:http://openpreservation.org/blog/2014/09/23/scape-project-ends-30th-september/>.

KUNY, Terry - The digital dark ages? Challenges in the preservation of electronic information. **International Preservation News**. [Em linha]17 (1998) 9–18. [Consult. 18 mai. 2015]. Disponível em WWW:<URL:http://www.ifla.org/files/assets/pac/ipn/17-98.pdf>. ISSN 08904960.

KURAMOTO, Hélio - Repositórios institucionais: políticas e mandatos. Em SAYÃO, LUIS *et al.* (Eds.) - **Implantação e gestão de repositório institucionais: políticas, memória, livre acesso e preservação** [Em linha]. Salvador : EDUFBA, 2009 Disponível em WWW:<URL:http://repositorio.ufba.br/ri/handle/ufba/473>. ISBN 978-85-232-0655-0. p. 203–217.

LAVOIE, Brian - The Open Archival Information System Reference Model: Introductory Guide. **DPC Technology Watch Series**. 2014). ISSN 2048-7916.

LAVOIE, Brian; DEMPSEY, Lorcan - Thirteen ways of looking at...digital preservation. **D-Lib Magazine**. [Em linha] 10:7–8 (2004). [Consult. 29 mar. 2014]. Disponível em WWW:<URL:http://www.dlib.org/dlib/july04/lavoie/07lavoie.html>. ISSN 10829873.

LEE, Kyong-ho *et al.* - The State of the Art and Practice in Digital Preservation. **J. Res. Natl. Inst. Stand. Technol.** [Em linha] 107:1 (2002) 93–106. [Consult. 18 mai. 2015]. Disponível em WWW:<URL:http://nvlpubs.nist.gov/nistpubs/jres/107/1/j71lee.pdf>.

LYMAN, Peter; BESSER, Howard - Defining the Problem of Our Vanishing Memory: background, current status, models for resolution. Em PARRY, ROSS (Ed.) - **Museums in a Digital Age**. 1. ed. Abrigdon : Routledge, 2010. ISBN 9780415402620. p. 336.

LYNCH, Clifford A. - Institutional Repositories: Essential Infraestructure for Scholarship in the Digital Age. **Association of Research Libraries: ARL**. [Em linha] 1:226 (2003) 1–16. [Consult. 3 mai. 2015]. Disponível em WWW:<URL:http://www.arl.org/storage/documents/publications/arl-br-226.pdf>. ISSN 1050-6098.

MARCONDES, Carlos Henrique; SAYÃO, Luis Fernando - Repositórios institucionais e livre acesso. Em SAYÃO, LUIS *et al.* (Eds.) - **Implantação e gestão de repositório institucionais: políticas, memória, livre acesso e preservação** [Em linha]. Salvador : EDUFBA, 2009 Disponível em WWW:<URL:http://repositorio.ufba.br/ri/handle/ufba/473>. ISBN 978-85-232-0655-0

MCGOVERN, Nancy Y. - **Digital preservation policy framework: Outline** [Em linha]. Michigan : ICPSR, 2007 (Relatório n.2). [Consult. 12 jan. 2017]. Disponível em WWW:<URL:http://www.icpsr.umich.edu/files/ICPSR/curation/preservation/policies/dp-policy-outline.pdf>.

MCHUGH, A. *et al.* - Risk management foundations for digital libraries: DRAMBORA (Digital Repository Audit Method Based on Risk Assessment). Em **Second Workshop on Foundations of Digital Libraries** [Em linha]. Budapest : University of Glasgow, 2007 [Consult. 14 mai. 2015]. Disponível em WWW:<URL:http://eprints.gla.ac.uk/33641/>.

MEDEIROS, Simone Assis; FERREIRA, Patrícia Aparecida - Política pública de acesso aberto à produção científica: um estudo sobre a implementação de repositórios institucionais em instituições de ensino superior. **Perspectivas em Gestão & Conhecimento**. [Em linha] 4:2 (2014) 195–217. [Consult. 9 mar. 2015]. Disponível em WWW:<URL:http://periodicos.ufpb.br/ojs2/index.php/pgc/article/view/16852>. ISSN 2236-417X.

MUELLER, Suzana Pinheiro Machado - A comunicação científica e o movimento de acesso livre ao conhecimento. **Ciência da Informação**. [Em linha] 35:2 (2006) 27–38. [Consult. 19 jun. 2014]. Disponível em WWW:<URL:http://revista.ibict.br/ciinf/index.php/ciinf/article/view/826/667>. ISSN 1518-8353.

NENTWICH, Michael - Project Outline and Background. Em **Cyberscience: Research in the Age of the Internet** [Em linha]. Vienna : Austrian Academy of Sciences Press, 2003 Disponível em WWW:<URL:http://hw.oeaw.ac.at/3188-7inhalt?frames=yes>. ISBN 3700131887. p. 3–20.

NENTWICH, Michael - Conceptual Framework: Definitions and a Model. Em **Cyberscience: Research in the Age of the Internet** [Em linha]. Vienna : Austrian

Academy of Sciences Press, 2003 [Consult. 11 jun. 2014]. Disponível em WWW:<URL:http://hw.oeaw.ac.at/3188-7inhalt?frames=yes>. ISBN 3700131887. p. 21–64.

NOONAN, Daniel W. - Digital Preservation Policy Framework : A Case Study. **Educause Review**. [Em linha]2014). [Consult. 27 jun. 2017]. Disponível em WWW:<URL:http://er.educause.edu/articles/2014/7/digital-preservation-policy-framework-a-case-study>. ISSN 1945-709X.

OPENAIRE - OPEN ACCESS INFRASTRUCTURE FOR RESEARCH IN EUROPE - **Open Access Policies and Mandates** [Em linha]. [S.l.] : OpenAIRE Project, 2015, atual. 2015. [Consult. 19 mai. 2015]. Disponível em WWW:<URL:https://www.openaire.eu/policies-and-mandates/open-access-pilot/open-access-policies-and-mandates>.

OPENDOAR - DIRECTORY OF OPEN ACCESS REPOSITORIES - **About OpenDOAR** [Em linha]. Nottingham : University of Nottingham, 2014 [Consult. 20 mai. 2015]. Disponível em WWW:<URL:http://www.opendoar.org/about.html>.

OSUL - **Digital Preservation Policy Framework** [Em linha]. Columbus : The Ohio State University Libraries, 2013 [Consult. 12 jan. 2017]. Disponível em WWW:<URL:https://library.osu.edu/documents/SDIWG/Digital_Preservation_Policy_Framework.pdf>.

PEARSON, David; POZO, Nick - **Explaining Pres Actions** [Em linha]. Canberra : National Library of Australia, 2009 (Relatório n.0.3.5). [Consult. 14 jul. 2015]. Disponível em WWW:<URL:http://pt.slideshare.net/natlibraryofaustralia/explaining-pres-actions>.

PENNOCK, Maureen - **Digital Preservation briefing paper** [Em linha]. Bath : Jisc, 2006 [Consult. 10 jun. 2014]. Disponível em WWW:<URL:http://www.jisc.ac.uk/publications/briefingpapers/2006/pub_digipreservationbp.aspx>.

PHILLIPS, M. *et al.* - **The NDSA levels of digital preservation: Explanation and Uses** [Em linha]. Washington, DC : Library of Congress, 2013 Disponível em WWW:<URL:http://www.digitalpreservation.gov/ndsa/working_groups/documents/ND

SA_Levels_Archiving_2013.pdf>.

PREMIS EDITORIAL COMMITTEE - **PREMIS Data Dictionary for Preservation Metadata** [Em linha]. 3. ed. Washington, D. C. : The Library of Congress, 2015 [Consult. 11 jul. 2016]. Disponível em WWW:<URL:http://www.loc.gov/standards/premis/v2/premis-dd-2-0.pdf>.

PRESERVICA - **Digital Preservation Maturity Model** [Em linha]. [S.l.] : PRESERVICA, 2014, atual. 2014. [Consult. 6 out. 2016]. Disponível em WWW:<URL:http://preservica.com/resource/praesent-ante-stiam-white-paper/>.

RAUCH, Carl; RAUBER, Andreas - Preserving Digital Media: Towards a Preservation Solution Evaluation Metric. Em **In Proceedings of the 7th International Conference on Asian Digital Libraries (ICADL 2004)** [Em linha]. Heidelberg : Springer Berlin Heidelberg, 2004 [Consult. 15 fev. 2015]. Disponível em WWW:<URL:http://ifs.tuwien.ac.at/dp/paper/Rau04_Utility_Analysis.pdf>.

REESE, Terry *et al.* - **Master Objects Repository Task Force Report** [Em linha]. Columbus : [s.n.] Disponível em WWW:<URL:https://library.osu.edu/document-registry/docs/401/download?1420058914>.

REPOSITÓRIUM - **FAQ's RepositóriUM** [Em linha]. Braga : Universidade do Minho, 2014, atual. 2014. [Consult. 28 jul. 2016]. Disponível em WWW:<URL:http://repositorium.sdum.uminho.pt/about/faqs/faqs.htm>.

RIBEIRO, Fanny Do Couto - **Análise de risco : uma metodologia a serviço da preservação digital** [Em linha]. Recife : Universidade Federal de Pernambuco, 2012 [Consult. 21 mai. 2015]. Disponível em WWW:<URL:http://www.ufpe.br/ppgci/images/documentos/disserta/2010 fanny.pdf>. Dissertação de Mestrado.

RICHARDSON, Roberto Jerry - **Pesquisa Social**. 3. ed. São Paulo : Atlas, 2015. ISBN 9788522421114.

RLG-NARA - RLG-NATIONAL ARCHIVES AND RECORDS ADMINISTRATION - **Trustworthy Repositories Audit & Certification: Criteria and Checklist** [Em linha]. Chicago : CRL-OCLC, 2007 Disponível em WWW:<URL:http://www.crl.edu/sites/default/files/attachments/pages/trac_0.pdf>.

RODRIGUES, Eloy - **Kit de Políticas Open Access** [Em linha]. Braga : Projeto RCAAP, 2009 [Consult. 30 jun. 2014]. Disponível em WWW:<URL:http://projeto.rcaap.pt/index.php/lang-pt/consultar-recursos-de-apoio/remository?func=startdown&id=336>.

ROLLEMBERG, Rodrigo - **Projeto de Lei do Senado Nº 387** [Em linha]. Brasília : Senado Federal, 2011 [Consult. 19 mai. 2015]. Disponível em WWW:<URL:http://www.senado.gov.br/atividade/materia/detalhes.asp?p_cod_mate=101006>.

RSP - REPOSITORIES SUPPORT PROJECT - **Policies & Legal Issues** [Em linha]. [S.l.] : JISC, 2013 [Consult. 19 mai. 2015]. Disponível em WWW:<URL:http://www.rsp.ac.uk/start/policies-and-legal-issues/>.

RUUSALEPP, Raivo *et al.* - Standards Alignment. Em **Aligning National Approaches to Digital Preservation** [Em linha]. Atlanta : Educopia Institute Publications, 2012 [Consult. 15 abr. 2015]. Disponível em WWW:<URL:http://www.cni.org/wp-content/uploads/2014/07/Aligning_National_Approaches_to_Digital_Preservation.pdf>. ISBN 978-0-9826653-1-2. p. 115–166.

RUUSALEPP, Raivo; DOBREVA, Milena - **Digital Preservation Services: State of the Art Analysis** [Em linha]. Roma : DC-NET Project, 2012, atual. 2012. [Consult. 30 jun. 2014]. Disponível em WWW:<URL:http://www.dc-net.org/getFile.php?id=467>.

SALZA, S. *et al.* - **Report on authenticity and plan for interoperable authenticity evaluation system** [Em linha]. Didcot : APARSEN, 2012 (Relatório n.D24.1). [Consult. 7 jul. 2016]. Disponível em WWW:<URL:http://www.alliancepermanentaccess.org/wp-content/plugins/download-monitor/download.php?id=79>.

SARACEVIC, Tefko - Ciência da informação: origem, evolução e relações. **Perspectivas em Ciência da Informação**. [Em linha] 1:1 (1996) 41–62. [Consult. 27 jul. 2016]. Disponível em WWW:<URL:http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/235>. ISSN 19815344.

SAYÃO, Luis *et al.* (EDS.) - **Implantação e gestão de repositórios institucionais:**

políticas, memória, livre acesso e preservação [Em linha]. Salvador : EDUFBA, 2009 [Consult. 25 abr. 2014]. Disponível em WWW:<URL:http://repositorio.ufba.br/ri/handle/ufba/473>. ISBN 978-85-232-0655-0.

SAYÃO, Luís Fernando - Interoperabilidade das bibliotecas digitais: o papel dos sistemas de identificadores persistentes - URN, PURL, DOI, Handle System, CrossRef e OpenURL. **Transinformação**. Campinas. [Em linha] 19:1 (2007) 65–82. [Consult. 16 fev. 2015]. Disponível em WWW:<URL:http://dx.doi.org/10.1590/S0103-37862007000100006>. ISSN 0103-3786.

SCAPE - **Project** [Em linha], atual. 2014. [Consult. 23 set. 2014]. Disponível em WWW:<URL:http://www.scape-project.eu/about/project>.

SDUM - SERVIÇOS DE DOCUMENTAÇÃO DA UNIVERSIDADE DO MINHO - **Sobre Repositórios OA** [Em linha]. Braga : Universidade do Minho, 2015 [Consult. 20 mai. 2015]. Disponível em WWW:<URL:http://openaccess.sdum.uminho.pt/?page_id=348>.

SEADLE, Michael - Archiving in the networked world: authenticity and integrity. **Library Hi Tech**. 2012). ISSN 0737-8831.

SHELDON, Madeline - **Analysis of Current Digital Preservation Policies** [Em linha]. [S.l.] : Library of Congress, 2013 [Consult. 9 jun. 2016]. Disponível em WWW:<URL:https://blogs.loc.gov/digitalpreservation/2013/08/analysis-of-current-digital-preservation-policies-archives-libraries-and-museums/>.

SHERPA-JULIET - **About JULIET** [Em linha]. [S.l.] : University of Nottingham, 2015, atual. 2015. [Consult. 19 mai. 2015]. Disponível em WWW:<URL:http://www.sherpa.ac.uk/juliet/index.php?la=en&mode=simple&page=about>.

SHIEBER, Stuart; SUBER, Peter - **Good practices for university open-access policies** [Em linha]. Cambridge : Havard Open Access Project, 2004, atual. 2004. [Consult. 19 mai. 2015]. Disponível em WWW:<URL:http://bit.ly/goodoa>.

SIERMAN, Barbara; JONES, Catherine; ELSTRØM, Gry - **Catalogue of Preservation Policy Elements** [Em linha]. Seibersdorf : SCAPE Project, 2014 (Relatório n.D13.2). Disponível em WWW:<URL:http://scape-project.eu/deliverable/d13-2-catalogue-of-

preservation-policy-elements>.

SMITH, MacKenzie *et al.* - DSpace: An Open Source Dynamic Digital Repository. **D-Lib Magazine**. [Em linha] 9:1 (2003). [Consult. 27 jul. 2016]. Disponível em WWW:<URL:http://www.dlib.org/dlib/january03/smith/01smith.html>. ISSN 1082-9873.

STRODL, Stephan *et al.* - How to choose a digital preservation strategy. Em **Proceedings of the ACM/IEEE Joint Conference on Digital Libraries (JCDL'07)** [Em linha]. Vancouver : ACM, 2007 [Consult. 9 fev. 2015]. Disponível em WWW:<URL:http://www.ifs.tuwien.ac.at/~becker/pubs/strodl_choose_JCDL07.pdf>.

SUBER, Peter - **Open Access Overview: focusing on open access to peer-reviewed research articles and their preprints** [Em linha]. Richmond : Earlham College, 2015, atual. 2015. [Consult. 20 mai. 2016]. Disponível em WWW:<URL:http://bit.ly/oa-overview>.

SUL - **Digital Preservation Policy for the State and University Library** [Em linha]. Aarhus : [s.n.] (Relatório n.4). [Consult. 4 out. 2016]. Disponível em WWW:<URL:http://en.statsbiblioteket.dk/About the library/DigitalPreservationPolicy.pdf>.

TANSLEY, R.; BASS, M.; SMITH, Mk - DSpace as an open archival information system: Current status and future directions. **Ecdl 2003**. [Em linha]2003) 446–460. [Consult. 27 jul. 2016]. Disponível em WWW:<URL:http://hdl.handle.net/1721.1/29464>. ISSN 0302-9743.

THIBODEAU, Kenneth - Overview of Technological Approaches to Digital Preservation and Challenges in Coming Years. Em **Proceedings of The State of Digital Preservation: An International Perspective** [Em linha]. Washington, DC : CLIR, 2002 [Consult. 18 mai. 2015]. Disponível em WWW:<URL:http://www.clir.org/pubs/reports/pub107/thibodeau.html>.

UFAL - **Repositório Institucional da UFAL** [Em linha], atual. 2016. [Consult. 18 mai. 2016]. Disponível em WWW:<URL:http://www.repositorio.ufal.br/>.

UFBA - **Sobre o RI UFBA: Documentos** [Em linha]. Salvador : UFBA, 2010, atual. 2010. [Consult. 28 jul. 2016]. Disponível em

WWW:<URL:https://repositorio.ufba.br/ri/about/about.jsp>.

UFC - **Política Institucional de Informação Técnico-Científica da UFC** [Em linha], atual. 2011. [Consult. 13 mai. 2016]. Disponível em WWW:<URL:http://www.repositorio.ufc.br/sobre/UFC_Consuni_2011_Resolucao02.pdf>.

UFES - **Repositório Institucional da UFES** [Em linha], atual. 2016. [Consult. 17 mai. 2016]. Disponível em WWW:<URL:http://repositorio.ufes.br/>.

UFF - **Repositório Institucional da UFF** [Em linha], atual. 2016. [Consult. 17 mai. 2016]. Disponível em WWW:<URL:http://www.repositorio.uff.br/jspui/>.

UFG - **Repositório da Universidade Federal de Goiás** [Em linha], atual. 2016. [Consult. 17 mai. 2016]. Disponível em WWW:<URL:http://repositorio.bc.ufg.br>.

UFGD - **Política Institucional para Divulgação das Produções Científicas e Técnicas da Universidade Federal da Grande Dourados** [Em linha] [Consult. 5 mai. 2016]. Disponível em WWW:<URL:http://www.ufgd.edu.br/prograd/repositorio-ufgd>.

UFJF - **Repositório Institucional da UFJF** [Em linha], atual. 2016. [Consult. 18 mai. 2016]. Disponível em WWW:<URL:https://repositorio.ufjf.br/jspui/>.

UFLA - **Política de Informação Institucional da UFLA** [Em linha], atual. 2012. [Consult. 13 mai. 2016]. Disponível em WWW:<URL:http://www.ufla.br/documentos/arquivos/082_13112012.pdf>.

UFMA - **Repositório Institucional da UFMA** [Em linha], atual. 2016. [Consult. 29 set. 2016]. Disponível em WWW:<URL:https://repositorio.ufma.br/jspui/>.

UFMG - **Repositório Digital da UFMG** [Em linha], atual. 2016. [Consult. 14 mai. 2016]. Disponível em WWW:<URL:https://dspaceprod02.grude.ufmg.br/dspace/>.

UFMS - **Repositório Institucional da UFMS** [Em linha], atual. 2016. [Consult. 17 mai. 2016]. Disponível em WWW:<URL:http://repositorio.cbc.ufms.br:8080/jspui/>.

UFMT - **Repositório Institucional da UFMT** [Em linha], atual. 2016. [Consult. 18 mai. 2016]. Disponível em WWW:<URL:http://200.129.241.94:8080/repositorio/>.

UFOP - **Política de Informação do Repositório Institucional da UFOP** [Em linha], atual. 2013. [Consult. 13 mai. 2016]. Disponível em

WWW:<URL:http://200.239.128.16/jspui-teste/image/resolucao_cepe.pdf>.

UFPA - **Repositório Institucional da UFPA** [Em linha], atual. 2016. [Consult. 17 mai. 2016]. Disponível em WWW:<URL:http://repositorio.ufpa.br/jspui/>.

UFPB - **Plano de Desenvolvimento Institucional 2014-2018** [Em linha]. João Pessoa : Universidade Federal da Paraíba, 2014 [Consult. 27 jul. 2016]. Disponível em WWW:<URL:http://www.proplan.ufpb.br/proplan/contents/documentos/pdi_ufpb_2014-2018.pdf>.

UFPB - **Política de Segurança da Informação** [Em linha]. João Pessoa : Conselho Universitário, 2014 [Consult. 27 jul. 2016]. Disponível em WWW:<URL:http://www.ufpb.br/cgti/contents/noticias/politica-de-seguranca-da-informacao-psi-na-ufpb/resolucao-32-2014.pdf>.

UFPB - **Repositório Institucional da UFPB** [Em linha], atual. 2016. [Consult. 17 mai. 2016]. Disponível em WWW:<URL:http://rei.biblioteca.ufpb.br/jspui/>.

UFPB - **BDTD - Orientações para Publicação** [Em linha]. João Pessoa : Biblioteca Central, 2016 [Consult. 27 jul. 2016]. Disponível em WWW:<URL:http://www.biblioteca.ufpb.br/biblioteca/contents/servicos/bdtd-orientacoes-para-publicacao>.

UFPE - **Repositório Institucional da UFPE** [Em linha], atual. 2016. [Consult. 17 mai. 2016]. Disponível em WWW:<URL:http://repositorio.ufpe.br/>.

UFPEL - **Política de Informação do Repositório Institucional Digital UFPel** [Em linha] [Consult. 5 mai. 2016]. Disponível em WWW:<URL:http://repositorio.ufpel.edu.br/handle/ri/2644>.

UFPI - **Repositório Institucional da UFPI** [Em linha], atual. 2016. [Consult. 18 mai. 2016]. Disponível em WWW:<URL:http://repositorio.ufpi.br/xmlui/>.

UFPR - **Repositório Institucional da UFPR** [Em linha], atual. 2016. [Consult. 17 mai. 2016]. Disponível em WWW:<URL:http://acervodigital.ufpr.br/>.

UFRB - **Política de Informação Técnico-Científica da UFRB** [Em linha], atual. 2013. [Consult. 13 mai. 2016]. Disponível em WWW:<URL:http://repositorio.ufrb.edu.br/arquivos/Portaria_771_2013.pdf>.

UFRGS - **Política Institucional para o Lume – Repositório Digital da UFRGS** [Em linha]. Porto Alegre : Gabinete do Reitor, 2010 [Consult. 27 jul. 2016]. Disponível em WWW:<URL:http://www.lume.ufrgs.br/arquivos_download/Portaria-5068.pdf>.

UFRN - **Política Institucional de Informação Técnico-Científica na Universidade Federal do Rio Grande do Norte** [Em linha] [Consult. 5 mai. 2016]. Disponível em WWW:<URL:http://repositorio.ufrn.br:8080/jspui/documentos/resolucao_592010_consepe_riufrn.pdf>.

UFS - **Política de acesso livre à informação científica da UFS** [Em linha] [Consult. 5 mai. 2016]. Disponível em WWW:<URL:https://ri.ufs.br/files/politica-ri-ufs.pdf>.

UFSC - **Repositório Institucional da UFSC** [Em linha], atual. 2016. [Consult. 17 mai. 2016]. Disponível em WWW:<URL:http://www.repositorio.ufsc.br/>.

UFSM - **Repositório Institucional da UFSM** [Em linha], atual. 2016. [Consult. 18 mai. 2016]. Disponível em WWW:<URL:http://repositorio.ufsm.br:8080/xmlui/>.

UFT - **Repositório Institucional da UFT: Orientações** [Em linha]. Palmas : SISBIB, 2011, atual. 2011. [Consult. 28 jul. 2016]. Disponível em WWW:<URL:https://repositorio.uft.edu.br/>.

UFU - **Repositório Institucional da UFU** [Em linha], atual. 2016. [Consult. 17 mai. 2016]. Disponível em WWW:<URL:http://repositorio.ufu.br/>.

UFV - **Repositório Institucional da UFV** [Em linha], atual. 2016. [Consult. 18 mai. 2016]. Disponível em WWW:<URL:http://www.locus.ufv.br/>.

UFVJM - **Política de Funcionamento do Repositório Institucional da UFVJM** [Em linha], atual. 2010. [Consult. 30 jun. 2016]. Disponível em WWW:<URL:http://acervo.ufvjm.edu.br/jspui/>.

UK DATA ARCHIVE - **Preservation policy** [Em linha]. Colchester : University of Essex, 2014 [Consult. 27 jul. 2016]. Disponível em WWW:<URL:http://data-archive.ac.uk/media/54776/ukda062-dps-preservationpolicy.pdf>.

UNB - **Política de Informação do Repositório da Universidade de Brasília** [Em linha], atual. 2013. [Consult. 13 mai. 2016]. Disponível em WWW:<URL:http://repositorio.unb.br/termo/resolucao.pdf>.

UNIFEI - **Repositório Institucional da UNIFEI** [Em linha], atual. 2016. [Consult. 29 set. 2016]. Disponível em WWW:<URL:https://repositorio.unifei.edu.br/xmlui/>.

UNIFESP - **Repositório Institucional da UNIFESP** [Em linha], atual. 2016. [Consult. 18 mai. 2016]. Disponível em WWW:<URL:http://repositorio.unifesp.br/>.

UNILA - **Repositório Institucional da UNILA** [Em linha], atual. 2013. [Consult. 20 jun. 2016]. Disponível em WWW:<URL:https://dspace.unila.edu.br/>.

UNIPAMPA - **Política de Informação Institucional da UNIPAMPA** [Em linha], atual. 2015. [Consult. 18 mai. 2016]. Disponível em WWW:<URL:http://porteiras.r.unipampa.edu.br/portais/sisbi/repositorio-digital/>.

UNIR - **Repositório Institucional da UNIR** [Em linha], atual. 2016. [Consult. 18 mai. 2016]. Disponível em WWW:<URL:http://www.ri.unir.br/jspui/>.

UTFPR - **Política de Informação do Repositório Institucional da UTFPR** [Em linha] [Consult. 5 mai. 2016]. Disponível em WWW:<URL:http://repositorio.utfpr.edu.br/jspui/sobre/politica_repositorio_1.pdf>.

WAUGH, Andrew *et al.* - Preserving digital information forever. Em **5th Conference on Digital Libraries** [Em linha]. New York : ACM, 2000 [Consult. 18 mai. 2015]. Disponível em WWW:<URL:http://dl.acm.org/citation.cfm?doid=336597.336659>. ISBN 158113231X

WEBB, Colin; PEARSON, David; KOERBIN, Paul - Oh, you wanted us to preserve that?! statements of preservation intent for the national library of Australia's digital collections. **D-Lib Magazine**. [Em linha] 19:1/2 (2013). [Consult. 18 fev. 2015]. Disponível em WWW:<URL:http://www.dlib.org/dlib/january13/webb/01webb.html>. ISSN 10829873.

WILSON, Andrew - **Significant Properties** [Em linha]. London : InSPECT Project, 2007 (Relatório n.2.2). [Consult. 18 mai. 2015]. Disponível em WWW:<URL:www.significantproperties.org.uk/wp22_significant_properties.pdf>.

YUL - **Digital Preservation Policy** [Em linha]. Yale : Yale Univeristy Library, 2015, atual. 2015. [Consult. 4 out. 2016]. Disponível em WWW:<URL:http://web.library.yale.edu/departments/preservation/policies-procedues-guidelines>.

ZIERAU, Maj-britt Olmütz - **A Holistic Approach to Bit Preservation** [Em linha]. Hvidrove : University of Copenhagen, 2011 [Consult. 18 mai. 2015]. Disponível em WWW:<URL:http://www.diku.dk/forskning/phd-studiet/phd/thesis_20111215.pdf>. Tese.

Anexo

A Preservação Digital nos Repositórios Institucionais das Universidades Federais.

Prezado(a) Administrador(a),

A preservação digital em repositórios institucionais é uma prática largamente difundida e apoiada em uma política de preservação digital em países da América do Norte e da Europa. No Brasil, apesar de existir idêntica preocupação com a preservação digital, ainda está em uma fase incipiente ou mesmo ausente nos repositórios institucionais. Nesse sentido, vimos solicitar a sua colaboração no preenchimento deste questionário que tem como objetivo coletar informações que possibilitem caracterizar a preservação digital nos Repositórios Institucionais das Universidades Federais. Se você concorda em participar desta pesquisa, por favor, responda cada item do questionário, selecionando a(s) opção(s) que mais se adequa(m) à sua condição de administrador do repositório fornecendo as demais informações solicitadas, quando se fizerem necessárias. Este formulário não grava o seu nome, entretanto se quiser obter uma cópia do resultado da pesquisa você poderá informar o seu *e-mail* ao final do questionário. De todo modo, por ocasião da publicação dos resultados da pesquisa nenhum dado de identificação será divulgado.

A pesquisa levará aproximadamente 10 minutos para ser respondida. Se tiver dúvidas ou questões, por favor entre em contato. Muito obrigado pela sua colaboração.

Laerte Pereira da Silva Júnior
skype: laertevoip
whatsapp: +351 919 403 590
e-mail: laerte.psjunior@gmail.com

Elementos para definição de uma política de preservação digital.

Esta seção visa ao levantamento das características gerais de uma política de preservação digital.

---

1. Os administradores das comunidades e coleções reclamam contra a falta de uma política de preservação digital?

- o Sim.
- o Não.
- o Não sei informar.

2. Quais são as políticas e normas institucionais que poderiam ser relacionadas com uma política de preservação digital?

- □ A Política Institucional de Informação do próprio repositório.
- □ A Política de Informação da Universidade.
- □ A Política de Desenvolvimento Institucional (PDI).
- □ A Política de Segurança da Informação da Universidade.
- □ Não sei informar.
- □ Outro: _____

3. Existem documentos que são de acesso restrito aos membros da sua Universidade?

- o Sim.
- o Não.

4. Que tipo de documento o repositório aceita para preservar? Por exemplo: teses, dissertações, artigos, videos de conferências, etc.

Resposta:

_____

5. O repositório estabelece um acordo com os depositantes para realizar uma conversão de formato de arquivo para fins de preservação e acesso, independentemente de um software específico?

- Sim.
- Não.

6. O repositório define critérios para exclusão de documentos ou de coleções?

- Sim, apenas para os documentos.
- Sim, apenas para as coleções.
- Sim, para ambos.
- Não.

7. O repositório possui um glossário de termos relacionados com a preservação digital?

- Sim.
- Não.

_____

Elementos para a implementação de uma política de preservação digital.

Esta seção visa ao levantamento das características de implementação, predominantemente técnicas, de uma política de preservação digital.

_____

8. A universidade possui um setor responsável pela preservação digital no repositório?

- o Sim.
- o Não.

9. A qual setor ou equipe está atribuída a responsabilidade pela preservação digital no repositório da sua universidade?

- o Não existe setor ou equipe com tal responsabilidade.
- o Biblioteca Central.
- o Superintendência (ou Núcleo) de Tecnologia da Informação.
- o Biblioteca Central e Superintendência (ou Núcleo) de Tecnologia da Informação.
- o Outros:

  _____

10. O pleno funcionamento de um repositório implica o investimento em infraestrutura tecnológica, capacitação da equipe tecnoadministrativa e preservação digital. A qual setor (ou setores) poderia ser atribuída a responsabilidade pelo planejamento financeiro que garanta o pleno funcionamento do repositório?

Resposta:

_____

11. Quais são os tipos de licença relacionada com os direitos autorais aplicados no ato da submissão de documentos?

- □ Declaração de distribuição não exclusiva.
- □ Licença *Creative Commons*.
- □ Nenhuma licença é aplicada.
- □ Outros: _____

12. No caso de o repositório aplicar algum tipo de licença, esta inclui alguma cláusula relativa à preservação do material que vai ser depositado?

- o Sim.
- o Não.

13. Assinale a posição da universidade face à inclusão do repositório na Rede Brasileira de Serviços de Preservação Digital coordenada pelo IBICT (Rede Cariniana).

- o O repositório está totalmente incluído na Rede Cariniana.
- o O repositório está parcialmente incluído na Rede Cariniana.
- o O repositório será incluído na Rede Cariniana.
- o Não se pretende incluir o repositório na Rede Cariniana.
- o Outro: _____

14. Quais são os padrões adotados pelo repositório?

- □ Dublin Core
- □ PREMIS
- □ METS
- □ Modelo de Referência OAIS
- □ Outro: _____

15. Quais são os tipos de formato de arquivo aceitos pelo repositório?

- o Formatos abertos.
- o Formatos proprietários.
- o Ambos.

16. Quem é o administrador geral do repositório?

- o Um professor.
- o Um bibliotecário.

- o Um servidor tecnoadministrativo.
- o Um grupo de bibliotecários.
- o Outro: _____

17. Quem valida os metadados dos objetos depositados no repositório?

- □ Bibliotecário.
- □ Técnico de suporte.
- □ Bolsista.
- □ Estagiário.
- □ Professor.
- □ Secretário de Departamento.
- □ Outro: _____

18. Quem atua no suporte e manutenção do repositório?

- □ Técnicos de informática da Biblioteca Central.
- □ Técnicos da Superintendência (ou Núcleo) de Tecnologia da Informação.
- □ Outro: _____

19. A equipe de administradores e técnicos do repositório tem participado de treinamentos ou cursos em preservação digital nos últimos 2 anos?

- o Sim.
- o Não.

20. A equipe de suporte técnico realiza os dois tipos de *backups* (o tradicional e o *Archival Information Package* – AIP) recomendados na documentação do *DSpace*?

- o Sim, para os dois tipos de *backups*.
- o Sim, apenas o *backup* tradicional.
- o Sim, apenas para o AIP.

- Não.
- Não sei informar.

21. A base de dados do repositório é rastreada com o antivírus recomendado pelo *DSpace*?

- Sim.
- Não.
- Não sei informar.

22. A funcionalidade de *checksum checker* do DSpace está sendo utilizada?

- Sim.
- Não.
- Não sei informar.

23. A funcionalidade para criar identificadores persistentes no *DSpace* usa o CNRI *Handle System*?

- Sim.
- Não.
- Não sei informar.

24. Existe um manual de preenchimento de metadados?

- Sim.
- Não.

Caso deseje receber o resultado da pesquisa, por favor, informe seu e-mail.

_____